Powerpoint Presentation Evaluation Based on Aggregation of Quality Criteria

Sergey Sakulin, Bauman Moscow State Technical University, Russia

https://orcid.org/0000-0001-9218-9725

Alexander Alfimtsev, Bauman Moscow State Technical University, Russia

https://orcid.org/0000-0002-3805-4499

Dmitry Sokolov, Bauman Moscow State Technical University, Russia

ABSTRACT

Today, there is no consensus about proper timing and conditions for integration of PowerPoint presentations into the educational process. But the model-based evaluation can make a decision-making process easier when it comes to using presentations. The purpose of this study is to build a formal model to evaluate presentations. In order to build a formal model, the authors suggest employing hierarchical structure consisting of aggregation operators, such as the weighted averaging operator, minimum operator, and fuzzy Choquet integral. The proposed formal model shows experts' knowledge of the interdependencies between the criteria. The experiment described in the paper demonstrates the effectiveness of such a model as it allows to formalize expert preferences gradually, which may include interdependencies between the quality criteria of a presentation. Thus, this model will allow to store, analyze, and compare presentations properties that are necessary for their successful application.

KEYWORDS

Aggregation Operator, Choquet Integral, Partial Data, Preference Relations, Presentation Evaluation, Quality Criteria

1. INTRODUCTION

MS PowerPoint (PP) as an instrument was initially designed for Macintosh in 1984 and then it was purchased by Microsoft (Gaskins, 2012). Over the years, the number of educators who use PowerPoint presentations (PPP) for teaching has become overwhelming because PP has been the most famous means of making up presentations. Indeed, PP is installed on more than 1 billion computers (Brock, 2017).

However, no consensus has been reached if presentations are worth employing in educational process.

On the one hand, (Fedisson & Braidic, 2007) describes an experiment dedicated to evaluation of PPP effectiveness. The results of this experiment prove that the use of PPP increases the students' level of comfort and, subsequently, improves their performance.

On the other hand, (Kedare et al., 2019) described an experiment where students who completed the course using the traditional teaching method (chalk, blackboard, lecture) showed better results compared to their counterparts who used PowerPoint presentations. This study confirms the idea that conventional lecture can be evaluated as more comprehensible and informative, since such lectures may create interactive learning environments in which pieces of slides are excluded (Cosgun, 2017).

Recent meta-analysis (Baker et al., 2018) has shown that there was hardly any difference in learning results when employing PPP or not. In (Kernbach et al., 2015) identifies some deterrent measures for PPP. In particular, full length sentences are hard to be put on the PPP slides to hold the sence (Farkas, 2009) and understanding (Yates & Orlikowski, 2007). Popular bullet lists foster an illusion of clarity and do not show the whole picture (Gabriel, 2008). Meanwhile, (Jenkins, 2012) studies the debates about the need to apply PPP.

Therefore, instead of contemplating on whether to use this tool, it appears that we should concentrate on how and when it should be used to help students in the learning process (Jourdan & Papp, 2013). Besides these two important issues, the issue of what kind of presentation it should be to help students in learning the best way is not the least important.

For example, PowerPoint can be very beneficial, but the material that is not pertinent to the lecture subject is harmful to students learning (Bartsch & Cobern, 2003). And the rise of form over content spoils the presentation (Grech, 2018). This corresponds to Mayer multimedia learning theory (Mayer, 2002). In (Hallewell & Crook, 2019) the lectures styles is analyzed and draws a conclusion that the presentation of the material should be consistent and coherent despite the individual styles of lecturers. Also the educators must use PowerPoint presentations in such a way that students understand that the presentation is supplementary to class attendance, not a replacement for it (Crawley & Frey, 2008). One way to tackle the issue is to create interactive PowerPoint presentations (Boyas, 2008).

There is a wide range of studies of PPP applications in various areas: physics (Erdemir, 2011), medicine (Grech, 2018); (Bamne & Bamne, 2016), accounting (Sugahara & Boland, 2006), astronomy (Miller & James, 2011), sociology (Hill et al., 2012), education of foreign language students (Gordani & Khajavi, 2019). Each of the disciplines needs PPP to meet its particular requirements to make students learn in the most effective way (Garrett, 2016).

To find out possible ways to improve PPP, the software engineers can be guided by scientific research carried out in various fields of activity, on the one hand, and by the students feedback (Williams et al., 2016) and students surveys (Cullen et al., 2018), (Szabo & Hastings, 2000) on the other hand.

Such surveys are usually carried out among students to find out their opinion about a certain PPP (Apperson et al., 2008). The results of such surveys are the PPP quality criteria values averaged over the students. That criteria correspond to various PPP qualities. In (Basturk, 2008), the author lists the following groups of quality criteria for a presentation: "content" and "design". These groups correspond to the same composite criteria, which, in turn, are the results of the convolution of criteria in the respective groups. To compare PPP and give preference to one of them, the less formalized experts evaluation is applied (Apperson et al., 2008, Akella, 2017, Bridges & Luks, 2016).

However, such an approach is constrained by the fact that it does not let some of the experts see clearly and in formal details the reasoning of the other experts. Moreover, the application of simple average for aggregation of certain PPP quality criteria, there is no chance to take into account their possible mutual influence that, in its turn, makes the evaluation of PPP inexact.

This brings us to the crying task of developing a formal approach to evaluate PPPs that would allow the identification of the experts reasoning and taking into account possible mutual influence of PPP quality criteria.

To make an educated decision backed up by experts' opinion, this paper suggests a formal evaluation of presentations that relies on quality criteria. The main principal is similar to the evaluation of web interfaces on the basis of user criteria (Alfimtsev et al., 2016). To develop such a model, it is necessary to set the criteria of presentation quality first and then, with the help of experts, set the criteria aggregation operator. After the model is developed, it can be used for the evaluation of a particular presentation.

This paper is organized as follows. Section 2 describes the proposed approach to developing a model for evaluating PPP quality. In particular, in subsection 2.1, the criteria for the quality of PPP are stated; in subsection 2.2, the proposed formal model for evaluating the presentation is shown; in subsection 2.3, a step-by-step procedure for evaluating the PPP is described. Section 3 contains the experiments and their results. In particular, in subsection 3.1, presentations selected for the experiments are described; in subsection 3.2, the proposed evaluation procedure is implemented; subsection 3.3 interprets the results. Section 4 presents conclusions drawn from the experiment results regarding the application of the proposed approach and suggests some directions for further research of this problem.

2. PROPOSED APPROACH

2.1. Criteria of Presentation Quality

Criteria for presentation quality can be set out in different ways. In (Apperson et al., 2008, Akella, 2017, Bridges & Luks, 2016), a number of similar criteria are discussed.

In particular, in (Bridges & Luks, 2016), possible reasoning of an expert regarding the content of the presentation is described. Formalization of such reasoning, we denote "quality of content" by the criterion G_1. In turn, it can be obtained as a result of aggregation of the following criteria: g_1^1—"breadth of considered problem", g_2^1—"focus on the problem", and g_3^1—"quality of the list of references".

Criteria $g_1^1 - g_3^1$ are evaluated by an expert by a 10-point scale. If there is no list of references or table of contents, the criteria g_2^1 and g_3^1 are impossible to evaluate. In this case, the content quality of such a presentation is also not evaluated. The criterion g_4^1 denotes the presence of the list of references, and g_5^1—the presence of the table of contents. These criteria are binary and take values from the set {0, 1}. The introduction of the criteria g_4^1 and g_5^1 implies the necessity of a table of contents and list of references in a good presentation.

The quality of the presentation also depends on its design. There are several quality criteria for design. One of such criteria is "the degree of structure of individual slides and the presentation in general" (Akella, 2017, Bridges & Luks, 2016). In (Apperson et al., 2008), the presence of sound in a presentation was determined by students to be significant (69% of respondents preferred a presentation with sound to the same one without it). Other design quality criteria are font parameters (size, character color, color combination (contrast; background; color choices); quality of animation, graphs, and tables. The above criteria will be merged into two abstract criteria: g_1^2—"design" (font parameters, color combination, animation quality) and g_2^2—"presentation and slide structure." The criteria g_1^2 and g_2^2 are also evaluated by an expert by a 10-point scale. As in the case of the criterion G_1 "quality of content," the lack of introduction in the presentation will make it impossible to evaluate the quality of the design due to its impact on the presentation structure. Therefore, the presence of the "presentation and slide structure" criterion g_2^2 makes sense only if the binary "presence of introduction" criterion g_3^2 is 1. Let G_2 denote the composite criterion "quality of design." In turn, G_2 will be resulting from the aggregation criteria $g_1^2 - g_3^2$.

Now we shall consider the mutual influence of the denoted criteria. The composite criteria G_1 "quality of content" and G_2 "quality of design" can be evaluated separately since the design and content of the same presentation can be done by different people, and it will not affect the quality of the presentation. Criteria within relevant independent G_1 and G_2 groups can influence each other. Consider the appropriate relation. The criteria g_1^1 "breadth of considered problem", g_2^1 "focus on topic", and g_3^1 "quality of the list of references" correlate positively with each other, as it is obvious that the breadth of considered problem directly depends on both list of references used and not straying from the topic. In addition, the criteria g_1^1 "breadth of considered problem" and g_2^1 "focus on topic" are interdependent and, in a sense, interchangeable. Meeting the criterion g_2^1 can have almost the same effect on the evaluation result as the meeting of both the criteria g_1^1 and g_2^1 due to the fact that if the content of the presentation corresponds to the topic, even if the designated problem is not fully considered, the presentation is most likely good from the informational standpoint. The criteria g_1^2 "design" and g_2^2 "presentation and slide structure" correlate positively - high-quality design can be achieved only if the presentation features a high-quality structure. In this case, it is logical to give preference to g_2^2 over g_1^2 because the structure is more important in terms of information perception than design when it comes to presentations. The influence of the binary criteria g_4^1, g_5^1 and g_3^2 on the other criteria is described earlier in this paper.

2.2. Formal Model for Presentation Evaluation

The arguments concerning the presentation quality criteria from the previous section can be formalized using the hierarchy of the corresponding aggregation operators. An aggregation operator is a function that depends on input criteria and features specified properties (Beliakov et al., 2016). The criteria themselves and the result of the aggregation are determined and take values from the unit interval [0, 1], and therefore, in the model of a formal quality evaluation for presentations, all the criteria should be reduced to the interval [0, 1]. They can be reduced by dividing the corresponding ratings on a 10-point scale by 10. In order to obtain the overall evaluation of the presentation quality Ω from the values of the initial criteria $g_1^1, \ldots, g_5^1$, $g_1^2, \ldots, g_3^2$, the aggregation operators should be chosen.

The selection of such operators is not an obvious task. The differences between the operators and the methods of setting out the corresponding hierarchies are due to the specifics of the tasks being solved. The examples of such application are integrated evaluation of student performance (Magadum & Bapat, 2018, Mane et al., 2014), engineering risk assessment (Rozhok & Tatarinov, 2019), evaluation of the technological process state (Sakulin et al., 2017), estimation of weak classifiers in pattern recognition (Samorodov, 2016) or evaluation in forensic medicine (Anderson et al., 2010).

The most common and relatively simple operators are minimum (min), maximum (max), and weighted averaging operators. However, using only them makes it impossible to take into account any interaction between the criteria, including correlation and interchangeability discussed earlier. In order to formalize expert knowledge of dependent criteria, discrete Choquet and Sugeno integrals are used (Samorodov, 2011). The main difference between the two is that the Choquet integral is based on linear operators, while the Sugeno integral is based on nonlinear operators (min and max). The Choquet integral is suitable for quantitative aggregation, while the Sugeno integral is more suitable for ordinal aggregation. The quality criteria of the presentation are quantitative, and the change in each of them should influence the result of the evaluation. Therefore, the Choquet integral was chosen as an aggregation operator for the dependent criteria.

In order to use the Choquet integral, the identification of fuzzy measure is needed. Fuzzy measure expresses the subjective weight or importance of each subset of criteria and can be defined as follows

(Grabisch et al., 1998): Fuzzy (discrete) measure is a function $\mu : \ 2^{J} \to [0,1]$, where 2^{J} is the set of all subsets of the criteria index set $J = \{1,..., \ H\}$, which satisfies the following conditions:

1) $\mu(\varnothing) = 0, \ \mu(J) = 1$;
2) $\forall D, B \subseteq J : \ D \subseteq B \Rightarrow \mu(D) \leq \mu(B)$.

The value of the $\mu(D)$ measure can be interpreted as "weight" or combined impact of the criteria included into the subset D of the criteria index set $J = \{1,..., \ H\}$. In the considered problem, the presentation quality criteria are the Choquet integral criteria.

The above considered interdependencies between the criteria can be formalized by the Choquet integral (Marichal, 2000).

Correlation is the best known of the dependencies between criteria. Two criteria $i, j \in J$ are positively (negatively) correlated if experts can observe a positive (negative) correlation between the contributions of two criteria to the aggregation result. Positive correlation will be expressed by the inequation $I(i, j) < 0$, but negative one – by $I(i, j) > 0$. In case the criteria i and j do not correlate, we come to the equation $I(i, j) = 0$.

Substitutiveness (complementarity) is another type of dependence. The idea of formalizing this type of dependencies using fuzzy measures was proposed by (Murofushi & Soneda, 1993). For example, if we have two criteria $i, j \in J$ and can suppose that the expert believes that meeting only one criterion causes almost the same effect as meeting of both. Here the importance of a pair of criteria is close to the importance of each of them individually, even if other criteria are present. In this case we see that the criteria i and j almost substitutive or interchangeable. The inequations like the ones for expressing positive and negative correlation criteria can be applied to express these dependencies by fuzzy measures and the Choquet integral.

The identification is complicated by the necessity to set a measure value for each subset of the criteria that shows experts preferences, which is impossible to implement in practice (Grabisch, 1997). Therefore, such identification is carried out using various indirect methods (Grabisch et al., 2008).

To use the Choquet integral, we have to identify a fuzzy measure on the basis of expert knowledge. This identification is complicated by exponential increasing complexity in the sense that it is necessary to set a value of fuzzy measure for each subset of criteria. Setting the values of all 2^{J} coefficients of the fuzzy measure $\mu(D), \ D \subseteq J$ is very difficult or even impossible for the expert. Note that even in case of three criteria to determine the fuzzy measure, it is necessary to obtain $2^{3} = 8$ coefficients. Despite this complexity the Choquet integral still can be applied in practice. For this Grabisch proposed the concept of κ - order fuzzy measure or κ - additive fuzzy measure (Grabisch et al., 2008). This order κ can be less than the number of aggregated criteria, $\kappa < |J| = H$. The essence of the κ - additivity concept lies in simplification of the task of fuzzy measures determination by excluding the dependencies between more than κ criteria. According to that, we chose three-additive Choquet integral. To identify the fuzzy measure for the model, we chose the method of least variance (Kojadinovic, 2007). This method allows us to obtain a unique fuzzy measure or none if it contradicts expert preferences; it also provides the objectivity as a result of aggregation in addition to the subjectivity of an expert's opinion.

Let's build up a hierarchy in the form of a tree of aggregation operators to evaluate the quality of presentations. The choice of operators the tree consists of will be based on the influence of the criteria on each other and on the result as a whole.

The composite criteria G_{1} "quality of content" and G_{2} "quality of design" are independent of each other. For that reason, we use a simple weighted averaging operator $AGG(G_{1}, G_{2}) = w_{1}G_{1} + w_{2}G_{2}$,

where w_1 and w_2 are weights determined by an expert directly or indirectly. The Fishburn's scheme is applicable to obtain the values w_1 and w_2 (Fishburn, 1970).

Due to the mentioned interchangeability of the criteria g_1^1 "breadth of considered problem" and g_2^1 "focus on topic," we can apply the Choquet integral for their aggregation. At the same time, it is necessary that we take into account the impact of the criteria g_4^1 "presence of the list of references" and g_5^1 "presence of introduction" on the composite criterion G_1 "quality of content" described earlier. To do this, we use the minimum operator, which will reduce the value of criterion G_1 to zero if at least one of the criteria g_4^1, g_5^1 takes a zero value. Since the criteria for g_1^2 "design" and g_2^2 "presentation and slide structure" are interdependent, the Choquet integral is applicable as an appropriate aggregation operator. Sometimes experts struggle to answer certain questions from the questionnaire. For that reason, we introduced the option "I find it difficult to answer this question". To cope with this, we will apply the Kolmogorov-Nagumo theorem in our evaluation model (Nagumo, 1930, Kolmogorov, 1930). According to this theorem, if there is no value for any criterion, it can be replaced by the average value of the remaining criteria, which will slightly affect the result (Dujmović & Tré, 2011).

The way we received the criterion G_1, we will obtain the criterion G_2 through the minimum operator, which will nullify the criterion value of criterion G_2 if there is no introduction in the presentation (nullification of the g_3^2 "presence of introduction" criterion). The resulting hierarchy of the aggregation operators is shown in Figure 1.

In accordance with the hierarchy above, the evaluation of the presentation quality Ω can be expressed with the formula:

$$\Omega = w_1 \min[g_4^1, g_5^1, C_{\mu_1}(g_1^1, g_2^1, g_3^1)] + w_2 \min[C_{\mu_2}(g_1^2, g_2^2), g_3^2] \tag{1}$$

Here, the Choquet integrals $C_{\mu_1}(g_1^1, g_2^1, g_3^1)$ and $C_{\mu_2}(g_1^2, g_2^2)$ together with the corresponding minimum operators are used to calculate the values of the composite criteria G_1 and G_2.

Figure 1. Hierarchy of aggregation operators
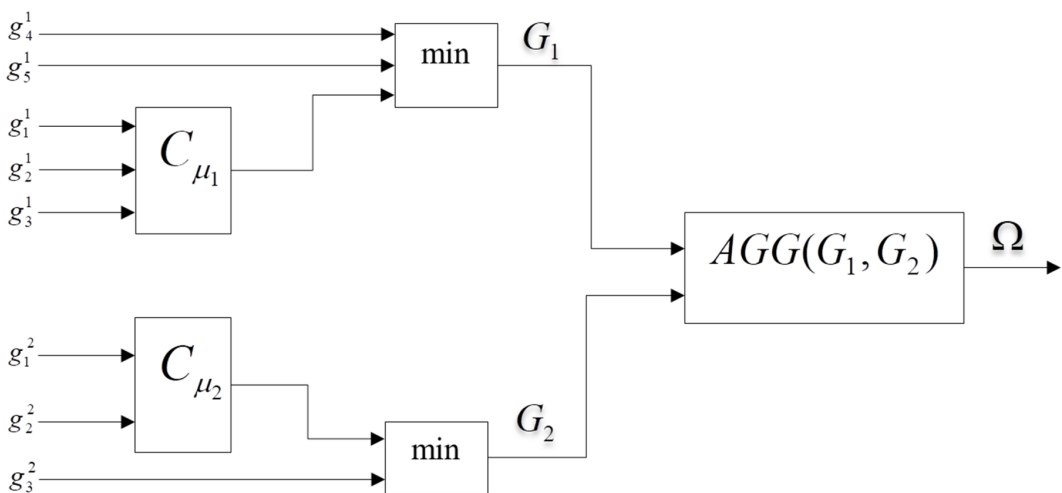
2.3. The Procedure for Assessing The Presentation Quality

Presentation evaluation based on the proposed model consists of the following steps.

Step 1. Provide experts with selected presentations and ask them to rate each presentation by the scale from 0 to 10 according the questionnaire:

1. Is the topic of the presentation fully considered?
2. How relevant is the content to the designated topic?
3. Evaluate the reliability of the sources.
4. Is there an introduction? (0 — no, 10 — yes)
5. Is there a list of references? (0 — no, 10 — yes)
6. Assess the design (graphics, sound elements, animation, color combination, font readability).
7. Assess the structure of the presentation and individual slides.
8. Is there a table of contents? (0 — no, 10 — yes)
9. Which is more important for the presentation, content or design? Or both?
10. Which presentation is better in terms of design?
11. Which presentation is better in terms of content?

If an expert finds it difficult to answer the question, they can leave the field in the questionnaire blank.

Step 2. Average the ratings obtained in the previous step on questions 1–8 by finding the arithmetic mean for each criterion. If there were any blanks in the questionnaire, the average of the existing ratings for the relevant criteria should be put there. The averaged ratings are divided by 10 for them to fit the interval [0,1].

The step should result in obtaining the values of the quality criteria $g_1^1, \ldots, g_5^1$, $g_1^2, \ldots, g_3^2$ for the presentations.

Step 3. Average the answers to question 9 by selecting the most common expert response. The result of processing the answers to this question will be presented as a ratio in the set {content, design}. For example, content ~ design means that, according to the experts, the presentation content is just as important as its design.

Answers to questions 10 and 11 should be averaged by choosing the best presentation based on the relevant experts ratings. The results of processing experts' answers to these questions will be presented as relationships of partial non-rigorous order between presentations. For example, $\mathbf{g}_1^1 \succ \mathbf{g}_2^1$, where $\succ$ means that, according to the experts, the presentation $\mathbf{g}_2$ is better than the presentation $\mathbf{g}_1$ in terms of content.

Step 4. Formalize the relationship between the criteria identified in Section 2.1 in the form of corresponding inequalities using the representation of the interaction of the Choquet integral.

Step 5. Apply the entropy maximization method to identify fuzzy measures μ_1, μ_2 based on the results from Steps 3 and 4 with the Kappalab package (Grabisch et al., 2008). Obtain the values of the weight coefficients w_1 and w_2 with the Fishburn weights scheme (Fishburn, 1970).

Step 6. Apply the values from the previous steps $g_1^1, \ldots, g_5^1$, $g_1^2, \ldots, g_3^2$ and fuzzy measures μ_1, μ_2 to formula (1) to obtain the result in the form of the overall quality evaluations Ω for each of the presentations using the Kappalab package.

3. EXPERIMENTS AND RESULTS

3.1. Presentations Used for The Experiments

The evaluation procedure was applied to two publicly available presentations. The first is an introductory presentation from the CSF course on basic machine learning concepts (Figure 2) by Stanford University (SU) (Re & Ma, 2019).

Figure 2. SU presentation on machine learning

Figure 3. MIT presentation on computer thinking and data science

The second presentation is an introductory lecture on computer thinking and data science by the Massachusetts Institute of Technology (MIT), which is devoted to optimization problems (Figure 3) (Guttag, 2019).

3.2. Implementation of The Procedure

We chose 30 first-year postgraduate students from the University as our experts.

Step 1. The experts receive the questionnaire and answer the questions.
Step 2. We averaged the obtained answers to questions 1–8. You may find the averaged results of the survey in the form of normalized criteria values in Table 1.

Both presentations have content, introduction, and a list of references (unit values in the table). The value of the criterion g_1^1 "breadth of considered problem" is the same for both presentations. The topics are explored to some extent in both presentations, but there are some flaws according to the experts. The main drawback of the MIT presentation is excess of information, while the SU presentation suffers from the lack of it. As for the criterion g_2^1 "focus on topic," the SU presentation has a higher rating as the information is very consistent with the topic throughout the presentation. The MIT presentation, on the other hand, has the information that is irrelevant to the main topic. The list of references in MIT presentation contains a list of scientific and educational literature. The one of the SU presentation includes the links to Internet resources that may not be reliable. Therefore, the value of the criterion g_3^1 "evaluation of the list of references" in the presentation of MIT is greater. From Table 1, it becomes clear that the criterion g_1^2 "design" in the MIT presentation is of a lesser value than the one in the SU presentation possibly due to an insufficient number of graphic elements

Table 1. Averaged results of experts ratings

Criterion	MIT (g_1)	SU (g_2)
g_1^1	0.8	0.8
g_2^1	0.8	0.9
g_3^1	0.9	0.7
g_4^1	1	1
g_5^1	1	1
g_1^2	0.7	0.8
g_2^2	0.7	0.8
g_3^2	1	1

(images, tables, graphs) in the MIT presentation. The value of the criterion g_2^2 "presentation and slide structure" is equal for both presentations.

Step 3. The answers to question 9 of the questionnaire revealed that most experts consider both design and content equally important for presentations. In our case, which is fairly simple, the weight coefficients w_1 and w_2 can be obtained with the Fishburn method as follows:

$$w_1 = w_2 = 0,5 \tag{2}$$

The answers to question 10 of the questionnaire show that the MIT presentation is superior to the SU one in terms of the criterion G_1 "quality of content". This experts preference is expressed as a relation between the corresponding implementations:

$$\mathbf{g}_1^1 \succ \mathbf{g}_2^1 \tag{3}$$

Here $\mathbf{g}_1^1$ denotes the implementation of the criteria $g_1^1, ..., g_5^1$ for the MIT presentation, and $\mathbf{g}_2^1$ — the implementation of the same criteria for the SU presentation. Based on the answers to question 11 of the questionnaire, we can conclude that, according to the majority of experts, the SU presentation has better design (G_2) than the MIT presentation. This expert preference is expressed as a relation between the corresponding implementations:

$$\mathbf{g}_1^2 \prec \mathbf{g}_2^2 \tag{4}$$

Here $\mathbf{g}_1^2$ denotes the implementation of the criteria $g_1^2, ..., g_3^2$ for the MIT presentation, and $\mathbf{g}_2^1$ — the implementation of the same criteria for the SU presentation.

Step 4. The interdependencies between the criteria from Section 2.1 are formalized. The positive correlation of criteria, $g_1^1, ..., g_3^1$, can be formalized using the sign of the interaction index of the relevant criteria:

$$I_{\mu_1}(1,2,3) < 0 \tag{5}$$

Here the index μ_1 indicates that the interaction index (5) refers to the integral $C_{\mu_1}(g_1^1, g_2^1, g_3^1)$. The positive correlation between the g_1^2 "design" and g_2^2 "presentation and slide structure" criteria is formalized using the sign of the interaction index of the relevant criteria:

$$I_{\mu_2}(1,2) < 0 \tag{6}$$

Here the index μ_2 indicates that the interaction index (6) refers to the integral $C_{\mu_2}(g_1^2, g_2^2)$. The experts reasoning states that, when determining the value of the criterion G_2 "quality of design", the criterion g_2^2 "presentation and slide structure" is more important than the criterion g_1^2 "design". Let us formalize it using the corresponding Shapley indices:

$$\Phi_{\mu_2}(1) < \Phi_{\mu_2}(2) \tag{7}$$

Step 5. Based on the results of Steps 3 and 4, the fuzzy measures μ_1 and μ_2 are identified as inequalities (2–7). For identification, we applied a method for minimizing the dispersion of fuzzy measures under given constraints, which is implemented in a specialized Kappalab package. In accordance with this method, preference relations (3, 4) are translated into inequalities:

$$C_{\mu_1}(0.8, 0.7, 0.9) - C_{\mu_1}(0.8, 0.8, 0.5) \geq \delta_c^{\mu_1} \tag{8}$$

$$C_{\mu_2}(0.9, 0.8) - C_{\mu_2}(0.7, 0.5) \geq \delta_c^{\mu_2} \tag{9}$$

Here $\delta_c^{\mu_1}$, $\delta_c^{\mu_2}$ are experts indifference thresholds. Criteria g_4^1, g_5^1 and g_3^2, when converting preference relations (3, 4) into inequalities (8, 9), are not considered as they have unit values and do not affect the results of aggregation in the form of values of the composite criteria G_1 "quality of content" and G_2 "quality of design".

Inequalities (5–7) are converted, respectively, into inequalities:

$$-1 \leq I(1, 2, 3) \leq -\delta_I^{\mu_1} \tag{10}$$

$$-1 \leq I(1, 2) \leq -\delta_I^{\mu_2} \tag{11}$$

$$\$(g_2^2) - \$(g_1^2) \geq \delta_\$^{\mu_2} \tag{12}$$

Here $\delta_I^{\mu_1}$, $\delta_I^{\mu_2}$, $\delta_\$^{\mu_2}$, $\delta_c^{\mu_1}$, $\delta_c^{\mu_2}$ are indifference thresholds set by the experts to identify relevant fuzzy measures. All input criteria are defined within {0, 0.1, ..., 0.9, 1}, while the restrictions on indifference threshold values are set to prevent selection of values for which the fuzzy measures identification problem does not have a solution. Based on this, the following values of these thresholds were set: $\delta_I^{\mu_1} = 0{,}005$, $\delta_I^{\mu_2} = 0{,}005$, $\delta_\$^{\mu_2} = 0{,}02$, $\delta_c^{\mu_1} = \delta_c^{\mu_2} = 0{,}055$.

The identification results in the form of the parameters of the corresponding integrals (interaction indices and Shapley indices) are shown in Table 2. The problem of weight coefficients w_1 and w_2 was addressed earlier (2).

In **Step 6**, based on the values of the criteria $g_1^1, ..., g_5^1$, $g_1^2, ..., g_3^2$ and the fuzzy measures μ_1, μ_2 in the form of the corresponding interaction indices and Shapley indices, the values of the Choquet integrals $C_{\mu_1}(g_1^1, g_2^1, g_3^1)$ and $C_{\mu_2}(g_1^2, g_2^2)$ for $\mathbf{g}_1$, $\mathbf{g}_2$ are obtained. These values are shown in Table 2 along with the values of the overall evaluation Ω for the reviewed presentations.

Let us suppose that, for some reason, we could not get from the experts the value of the criterion g_2^1 "focus on topic." In this case, in accordance with the Kolmogorov-Nagumo theorem, the value of this criterion is replaced with the average value of the remaining criteria for each presentation. The results of aggregation for such a replacement are shown in Table 3.

3.3. Interpretation of The Results

The results obtained during the experiments are shown in Table 2. The interaction index $I_{\mu_1}(1,2)$ takes the value -0.0050 from the interval $[-1; -\delta_I^{\mu_1}]$. This expresses the interdependence of the criteria g_1^1, g_2^1, which corresponds to the experts reasoning. The remaining interaction indices for the criteria related to $C_{\mu_1}(g_1^1, g_2^1, g_3^1)$ are equal to 0. As expected, the results of the aggregation reflect the experts opinion formulated in (1–7). The obtained values of the Choquet integrals $C_{\mu_1}(g_1^1, g_2^1, g_3^1)$ and $C_{\mu_2}(g_1^2, g_2^2)$, given in the table, demonstrate the experts preference of the MIT presentation in terms of the criterion G_1 "quality of content" and the SU presentation in

Table 2. The results of criteria aggregation

	$\Phi(1)$	$\Phi(2)$	$\Phi(3)$	$I_{\mu_1}(1,2,3)$	$I_{\mu_2}(1,2)$	$\mathbf{g}_1$	$\mathbf{g}_2$
C_{μ_1}	0.4975	0.4975	0.005	-0.005	–	0.799	0.7969997
C_{μ_2}	0.5025	0.5025	–	–	-0.005	0.7	0.85025
Ω	–	–	–	–	–	0.7495	0.8236248

Table 3. The results of criteria aggregation with partial absence of input data

	$\mathbf{g}_1$	$\mathbf{g}_2$
C_{μ_1}	0.8003502	0.7969998
C_{μ_2}	0.7	0.85025
Ω	0.7501751	0.8236249

terms of the criterion G_2 "quality of design." The results of the aggregation also reflect the experts opinion that the higher the quality of the material presented (criteria g_1^1 and g_2^1), the lower the importance of the list of references evaluation (criterion g_3^1). Also, as expected, the evaluation of the composite criterion G_1 "quality of content" is higher for the implementation $\mathbf{g}_2$. Overall, the SU presentation is evaluated higher than the MIT presentation, which is shown in the values of the final evaluation Ω.

The results in the absence of information from the experts on a g_2^1 criterion value are shown in Table 3. The overall evaluation of the SU presentation remained relatively higher than that of the MIT presentation. The difference in final evaluation Ω has changed by 4.9%. This confirms that the proposed model remains successful and healthy in the partial absence of data from experts.

Thus, all the experts reasoning about the quality criteria of preferences were taken care of for the formal model of presentation evaluation.

4. DISCUSSION

Both technologies in educational process and technological literacy of lecturers on their own do not always mean their effective application. Their integration and the effective use require lecturers' knowledge and understanding of the theoretical basis that would justify and support the application of technologies (Yang & Skelcher, 2019).

Not only the theoretical basis for the application of technologies, but strong beliefs and long-held stereotypes of academic staff influence the application of the technologies (Tondeur e al., 2017). Moreover, it is not only important to introduce the technologies, but to use them in an intelligent way. The reasonable decision to use a technology in the educational process, informal or weakly formalized methods of analysis are employed. For example, (Page, 2019) compares and contrasts the use of tactile and digital sketches when designing, (Costa, Teixeira & Alvelos, 2018) analyses the current application of wholescale free online courses at universities, (Tafazoli, María & Abril, 2019) analyses the conditions of the successful introduction of intelligent computer-assisted language learning, (Tuapawa, 2017) brings out some features of educational online technologies application.

A recent tendency to converge various fields of expertise has not spared such fields of science as pedagogy and education, on the one hand, and formal theory of decision making, on the other hand. Therefore, there are formal approaches to decision making in the field of education. In particular, (Deb et al., 2019) discusses the application of fuzzy logic to find possible gaps in teaching students; (Singh & Rawani, 2019) discusses the fuzzy approach to rate various expectations of students of technical universities; (Yassine, Battou & Omar, 2019) describes how to use fuzzy analytical hierarchy process to define the weights of open online courses to increase completion rates and to decrease dropout rates; (Bendjebar, Lafifi & Seridi, 2016) employs fuzzy logic for a formal estimation of tutor profiles to classify online tutors. The results of the practical application of such formal approaches have shown the effectiveness and repeatability of the decisions made.

At the same time, the application of such formal approaches faces the challenges that involve a lack of means to formalize possible interconnections between the criteria to make a decision as well as possible lack of data. Methods of decision-making theory in possible accounting for interconnections between the criteria has allowed us to formalize experts' preferences with high accuracy and 'fine tuning' of the model. One of the most popular application of technologies in educational process is the use of Power Point. Therefore the model was used to evaluate PP presentations. The model takes into account the interconnections between the criteria and nonsensitive to a lack of data.

5. CONCLUSION AND RESEARCH PROSPECTS

Recent achievements in the theory of decision making have allowed us to formalize human reasoning where there is uncertainty and (or) some data is missed. The formal evaluation of PPP is bridged by such research areas as pedagogy, psychology and computer sciences.

The paper considers a model of PPP evaluation on the basis of hierarchy of aggregation operators including the Choquet integrals. The hierarchy is tolerant to the absence of input data due to the application of the Kolmogorov-Nagumo approach. This model allows PPP evaluation even if a part of experts' evaluations is missed and takes into account the interdependencies between the criteria as well. The results of the experiments on PPP evaluation have showed that all the experts' preferences were taken into account by the model built by the suggested approach.

In addition, such an approach faces a number of challenges such as choice of certain quality criteria of PPP, type of hierarchy for aggregation operators, choice of aggregation operators themselves, parameters of the aggregation operators. All this depends on the personality of the expert in a certain applied field of activity, on their subjective opinion.

Moreover, poor intuitive understanding of such aggregation operators as the Choquet integral makes experts employ simplified models that do not allow formalization of interdependencies between quality criteria.

Nevertheless, by increasing refinement it becomes possible to develop formal models for PPP evaluation in various fields of activity. If the construction of the described hierarchy is performed by several experts who may have different opinions about the subject, the use of the considered approach will help them find a compromise solution. This is possible because all considered types of expert preferences can be presented explicitly and unambiguously in the form of the corresponding operators incorporated in the hierarchy. The process of formalizing presentation evaluation can be iterative, which implies a gradual refinement of experts preferences. After such models are developed, it will become possible to store the data about the necessary properties of PPP in various fields of activity. This in its turn will allow to identify the differences more carefully between PPP properties applying them to certain areas of activity and provide essential recommendations for PPP designers.

The research plans in this area include developing software with deep learning to improve the experience of presentation evaluation as well as to building automatic pattern recognition models to evaluate PPP. Besides, to overcome the difficulties the experts face when working with aggregation operators, additional research is needed in the field of visualization of aggregation operators by cognitive graphics to help the experts in intuitive understanding.

REFERENCES

Akella, D. (2017). *Powerpoint vs. Recorded Lectures: An African American Perspective on Online Education*. IGI Global. doi:10.4018/978-1-5225-0968-4.ch013

Alfimtsev, A., Sakulin, S., & Levanov, A. (2016). Formalization of expert knowledge about the usability of web pages based on user criteria aggregation. *International Journal of Software Innovation*, *4*(3), 38–50. doi:10.4018/IJSI.2016070103

Anderson, M. F., Anderson, D. T., & Wescott, D. J. (2010). Estimation of adult skeletal age at death using the Sugeno fuzzy integral. *American Journal of Physical Anthropology*, *142*(1), 30–41. PMID:19845027

Apperson, J. M., Laws, E. L., & Scepansky, J. A. (2008). An assessment of student preferences for PowerPoint presentation structure in undergraduate courses. *Computers & Education*, *50*(1), 148–153. doi:10.1016/j.compedu.2006.04.003

Baker, J. P., Goodboy, A. K., Bowman, N. D., & Wright, A. A. (2018). Does teaching with PowerPoint increase students' learning? A meta-analysis. *Computers & Education*, *126*, 376–387. doi:10.1016/j.compedu.2018.08.003

Bamne, S. N., & Bamne, A. S. (2016). Comparative study of chalkboard teaching over PowerPoint teaching as a teaching tool in undergraduate medical teaching. *International Journal of Medical Science and Public Health*, *5*(12), 2585–2587. doi:10.5455/ijmsph.2016.01072016532

Bartsch, R. A., & Cobern, K. M. (2003). Effectiveness of PowerPoint presentations in lectures. *Computers & Education*, *41*(1), 77–86. doi:10.1016/S0360-1315(03)00027-7

Basturk, R. (2008). Applying the many facet Rasch model to evaluate PowerPoint presentation performance in higher education. *Assessment & Evaluation in Higher Education*, *33*(4), 431–444. doi:10.1080/02602930701562775

Beliakov, G., Sola, H. B., & Sánchez, T. C. (2016). *A practical guide to averaging functions*. Springer. doi:10.1007/978-3-319-24753-3

Bendjebar, S., Lafifi, Y., & Seridi, H. (2016). Modeling and Evaluating Tutors' Function using Data Mining and Fuzzy Logic Techniques. *International Journal of Web-Based Learning and Teaching Technologies*, *11*(2), 39–60. doi:10.4018/IJWLTT.2016040103

Boyas, E. A. (2008). Using PowerPoint to Encourage Active Learning: A Tool to Enhance Student Learning in the First Accounting Course. *International Journal of Information and Communication Technology Education*, *4*(2), 14–25. doi:10.4018/jicte.2008040102

Bridges, T. M., & Luks, A. M. (2016). *How to Give a Great PowerPoint Presentation*. Springer. doi:10.1007/978-3-319-33193-5_8

Brock, D. C. (2017). The improbable origins of Powerpoint. *IEEE Spectrum*, *54*(11), 42–49. doi:10.1109/MSPEC.2017.8093800

Cosgun Ögeyik, M. (2017). The effectiveness of PowerPoint presentation and conventional lecture on pedagogical content knowledge attainment. *Innovations in Education and Teaching International*, *54*(5), 503–510. doi:10.1080/14703297.2016.1250663

Costa, C., Teixeira, L., & Alvelos, H. (2018). Exploring the Usage of MOOCs in Higher Education Institutions: Characterization of the Most Used Platforms. *International Journal of Information and Communication Technology Education*, *14*(4), 1–17. doi:10.4018/IJICTE.2018100101

Crawley, D. C., & Frey, B. A. (2008). Examining the relationship between course management systems, presentation software, and student learning: An exploratory factor analysis. *International Journal of Information and Communication Technology Education*, *4*(1), 1–14. doi:10.4018/jicte.2008010101

Cullen, A. E., Williams, J. L., & McCarley, N. G. (2018). Conscientiousness and Learning via Feedback to Identify Relevant Information on PowerPoint Slides. *North American Journal of Psychology*, *20*(2), 425–443.

Deb, K., Banerjee, S., Das, A., & Bag, R. (2019). Tutorial Gap Identification Towards Student Modeling Using Fuzzy Logic. *International Journal of Information and Communication Technology Education*, *15*(3), 30–41. doi:10.4018/IJICTE.2019070103

Dujmović, J., & Tré, G. D. (2011). Multicriteria methods and logic aggregation in suitability maps. *International Journal of Intelligent Systems*, *26*(10), 971–1001. doi:10.1002/int.20509

Erdemir, N. (2011). The effect of PowerPoint and traditional lectures on students' achievement in physics. *Journal of Turkish Science Education*, *8*, 176–189.

Farkas, D. K. (2009). Managing three mediation effects that influence PowerPoint deck authoring. *Technical Communication (Washington)*, *56*, 1–11.

Fedisson, M., & Braidic, S. (2007). PowerPoint presentations increase achievement and student attitudes towards technology. *International Journal of Information and Communication Technology Education*, *3*(4), 64–75. doi:10.4018/jicte.2007100106

Fishburn, P. C. (1970). *Utility theory for decision making*. Wiley. doi:10.21236/AD0708563

Gabriel, Y. (2008). Against the tyranny of PowerPoint: Technology-in-use and technology abuse. *Organization Studies*, *29*(2), 255–276. doi:10.1177/0170840607079536

Garrett, N. (2016). How Do Academic Disciplines Use PowerPoint? *Innovative Higher Education*, *41*(5), 365–380. doi:10.1007/s10755-016-9381-8

Gaskins, R. (2012). *Sweating bullets: Notes about inventing PowerPoint*. Vinland Books.

Gordani, Y., & Khajavi, Y. (2019). The impacts of multi-modal PowerPoint presentation on the EFL students' content knowledge attainment and retention over time. *Education and Information Technologies*, 1–15.

Grabisch, M. (1997). K-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, *92*(2), 167–189. doi:10.1016/S0165-0114(97)00168-1

Grabisch, M., Kojadinovic, I., & Meyer, P. (2008). A review of methods for capacity identification in Choquet integral based multi-attribute utility theory: Applications of the Kappalab R package. *European Journal of Operational Research*, *186*(2), 766–785. doi:10.1016/j.ejor.2007.02.025

Grabisch, M., Orlovski, S. A., & Yager, R. R. (1998). *Fuzzy aggregation of numerical preferences*. Springer. doi:10.1007/978-1-4615-5645-9_2

Grech, V. (2018). The application of the Mayer multimedia learning theory to medical PowerPoint slide show presentations. *Journal of Visual Communication in Medicine*, *41*(1), 36–41. doi:10.1080/17453054.2017.1408400 PMID:29381105

Grech, V. (2018). WASP (Write a scientific paper): Optimisation of PowerPoint presentations and skills. *Early Human Development*, *125*, 53–56. doi:10.1016/j.earlhumdev.2018.06.006 PMID:29929910

Guttag, J. (2019). *Introduction and Optimization Problems*. Massachusetts Institute of Technology. Retrieved from https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-0002-introduction-to-computational-thinking-and-data-science-fall-2016/lecture-slides-and-files/MIT6_0002F16_lec1.pdf

Hallewell, M. J., & Crook, C. (2019). Performing PowerPoint lectures: Examining the extent of slide-text integration into lecturers' spoken expositions. *Journal of Further and Higher Education*, 1–16. doi:10.1080/0309877X.2019.1579895

Hill, A., Arford, T., Lubitow, A., & Smollin, L. M. (2012). "I'm Ambivalent about It" The Dilemmas of PowerPoint. *Teaching Sociology*, *40*(3), 242–256. doi:10.1177/0092055X12444071

Jenkins, E. (2012). Death by PowerPoint? A Media Ecology Examination of the Debates Over Slideware. *Explorations in Media Ecology*, *10*(3-4), 225–245. doi:10.1386/eme.10.3-4.225_1

Jourdan, L., & Papp, R. (2013). PowerPoint: It's not "Yes" or "No"–it's "When" and "How". *Research in Higher Education*, *22*, 1–12.

Kedare, R. V., Kharat, R. D., & Wagh, R. J. (2019). Impact of PowerPoint and Chalkboard teaching in Physiotherapy Undergraduates. *International Journal of Clinical and Biomedical Research*, *5*(1), 9–11. doi:10.31878/ijcbr.2018.51.03

Kernbach, S., Bresciani, S., & Eppler, M. J. (2015). Slip-sliding-away: A review of the literature on the constraining qualities of PowerPoint. *Business and Professional Communication Quarterly*, *78*(3), 292–313. doi:10.1177/2329490615595499

Kojadinovic, I. (2007). Minimum variance capacity identification. *European Journal of Operational Research*, *177*(1), 498–514. doi:10.1016/j.ejor.2005.10.059

Kolmogorov, A. N. (1930). Sur la notion de la moyenne. *Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez.*, *12*, 388–391.

Magadum, C. G., & Bapat, M. S. (2018). Ranking of Students for Admission Process by Using Choquet Integral. *International Journal of Fuzzy Mathematical Archive*, *15*(2), 105–113.

Mane, A. M., Dongale, T. D., & Bapat, M. S. (2014). Application of fuzzy measure and fuzzy integral in students failure decision making. *IOSR Journal of Mathematics*, *10*(6), 47–53. doi:10.9790/5728-10634753

Marichal, J.-L. (2000). An axiomatic approach to the discrete Choquet integral as a tool to aggregate interacting criteria. *IEEE Transactions on Fuzzy Systems*, *8*(6), 800–807. doi:10.1109/91.890347

Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, *41*, 85–139. doi:10.1016/S0079-7421(02)80005-6

Miller, S. T., & James, C. R. (2011). The effect of animations within PowerPoint presentations on learning introductory astronomy. *Astronomy Education Review*, *10*(1), 1–13. doi:10.3847/AER2010041

Murofushi, T., & Soneda, S. (1993). Techniques for reading fuzzy measures (III): interaction index. *9th Fuzzy System Symposium*, 693–696.

Nagumo, M. (1930). Über eine klasse der mittelwerte. *Japanese Journal of Mathematics: Transactions and Abstracts*, *7*, 71-79.

Page, T. (2019). A Comparison of Haptic Sketching and Digital Sketching: Considerations of Final Year Design Students. *International Journal of Information and Communication Technology Education*, *15*(2), 146–161. doi:10.4018/IJICTE.2019040109

Re, C., & Ma, T. (2019). *Machine Learning*. Stanford University. Retrieved from http://cs229.stanford.edu/notes-spring2019/lecture1_slide.pdf

Rozhok, A., & Tatarinov, V. (2019). The application of dynamic risk analysis methods for safety improvement on highways located near or on the territory of hydraulic engineering dams. *IOP Conference Series. Materials Science and Engineering*, *492*(1), 1–6. doi:10.1088/1757-899X/492/1/012005

Sakulin, S. A., & Alfimtsev, A. N. (2017). *Data fusion based on the fuzzy integral: model, methods and applications*. Nova Science Publishers.

Samorodov, A. V. (2011). Application of a fuzzy integral for weak classifiers boosting. *Pattern Recognition and Image Analysis*, *21*(2), 206–210. doi:10.1134/S1054661811020945

Samorodov, A. V. (2016). Method of weak classifiers fuzzy boosting: Iterative learning of quasi-linear algorithmic composition. *Pattern Recognition and Image Analysis*, *26*(2), 266–273. doi:10.1134/S105466181602019X

Singh, A. K., & Rawani, A. M. (2019). A Fuzzy Approach for Ranking of Student's Expectation From a Technical Institute. *International Journal of Smart Education and Urban Society*, *10*(2), 41–52. doi:10.4018/IJSEUS.2019040103

Sugahara, S., & Boland, G. (2006). The effectiveness of PowerPoint presentations in the accounting classroom. *Accounting Education*, *15*(4), 391–403. doi:10.1080/09639280601011099

Szabo, A., & Hastings, N. (2000). Using IT in the undergraduate classroom: Should we replace the blackboard with PowerPoint? *Computers & Education*, *35*(3), 175–187. doi:10.1016/S0360-1315(00)00030-0

Tafazoli, D., María, E. G., & Abril, C. A. H. (2019). Intelligent Language Tutoring System: Integrating Intelligent Computer-Assisted Language Learning Into Language Education. *International Journal of Information and Communication Technology Education*, *15*(3), 60–74. doi:10.4018/IJICTE.2019070105

Tondeur, J., Van Braak, J., Ertmer, P. A., & Ottenbreit-Leftwich, A. (2017). Understanding the relationship between teachers' pedagogical beliefs and technology use in education: A systematic review of qualitative evidence. *Educational Technology Research and Development*, *65*(3), 555–575. doi:10.1007/s11423-016-9481-2

Tuapawa, K. (2017). Educational online technologies in blended tertiary environments: Experts' perspectives. *International Journal of Information and Communication Technology Education*, *13*(3), 1–14. doi:10.4018/IJICTE.2017070101

Williams, J. L., McCarley, N. G., Haynes, J. M., Williams, E. H., Whetzel, T., Reilly, T., & Bailey, L. et al. (2016). The Use of Feedback to Help College Students Identify Relevant Information on PowerPoint Slides. *North American Journal of Psychology*, *18*(2), 239–256.

Yang, D., & Skelcher, S. (2019). Improving Teachers' Understanding of Theoretical Foundations of Technology Use: Connecting Theories With Technology. *International Journal of Information and Communication Technology Education*, *15*(4), 113–127. doi:10.4018/IJICTE.2019100108

Yassine, A. F. A., Battou, A., & Omar, B. A. Z. (2019). Learning Through Massive Open Online Courses Platforms Based on Fuzzy Analytic Hierarchy Process. *International Journal of Smart Education and Urban Society*, *10*(3), 1–12. doi:10.4018/IJSEUS.2019070101

Yates, J., & Orlikowski, W. (2007). The PowerPoint presentation and its corollaries: How genres shape communicative action in organizations. *Communicative practices in workplaces and the professions: Cultural perspectives on the regulation of discourse and organizations*, *1*, 67-92.

Sergey Sakulin graduated the Bauman Moscow State Technical University in 2001. He is a Ph.D. by BMSTU in 2009. Today he is assistant professor of Information systems and telecommunications department. He has thirty scientific papers and five educational books. Scientific interests lie in the fields of artificial intelligence methods and expert knowledge formalization and visualization.

Alexander Alfimtsev received the B. S. and M. S. degrees in computer science from the Bauman Moscow State Technical University (BMSTU), Russia, in 2003 and 2005 respectively, and the Ph. D. degree in 2008. He is currently a Professor of Information systems and telecommunications department at BMSTU. He is also a Member of the International Institute of Informatics and Systemics, USA. He has 70 scientific papers, including 5 patents for inventions. His main research interests include the use of computer vision and artificial intelligence methods for human-computer interaction, pattern recognition and multimodal interfaces.

Dmitry Sokolov is an MSc student at Bauman Moscow State Technical University. He has five scientific papers. Scientific interests lie in the fields of artificial intelligence methods and programming engineering.