Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

Issues of organization of expertise and problems of expert assessments

To cite this article: P I Paderno *et al* 2020 *J. Phys.: Conf. Ser.* **1703** 012047

Issues of organization of expertise and problems of expert assessments

P I Paderno[1], **E A Burkov**[2] **and E A Lavrov**[3]

[1] Department of Information Systems, Saint Petersburg Electrotechnical University «LETI», St. Petersburg, Professora Popova 5, 197376, Russia
[2] Department of Algorithmic Mathematics, Saint Petersburg Electrotechnical University «LETI», St. Petersburg, Professora Popova 5, 197376, Russia
[3] Department of Computer Science, Sumy State University, Sumy, Rymskogo-Korsakova 2, 40007, Ukraine

E-mail: eaburkov@gmail.com

Abstract. In the course of preparation and carrying out of various expertises, in a number of cases methodological and organizational problems arise due to both the insufficiently high level of training of the working group members who develop the program and methodology of the expertise, and various restrictions, that could have a material impact on the credibility of the results. These restrictions generally are either financial in nature or are associated with a lack of sufficiently qualified and objective experts in the subject area. Objective: to develop a set of techniques that allow to expand the possibilities of using expert assessment methods in the preparation and carrying out of modern expertise in the presence of restrictions associated with external conditions. Methods: methods and approaches of system analysis, integration of individual expert assessments are used. Results: a comprehensive analysis of the problems faced by specialists in the organization of various expertises has been carried out. A number of techniques have been proposed that make it possible to significantly simplify the process of preparing and carrying out expertises but at the same time do not negatively affect the adequacy of the results obtained. Practical significance: the results obtained can be widely used in the course of various expertises, both in the process of selecting experts and in the process of direct obtaining and combining expert assessments. In addition, the use of obtained results in terms of different organizational and material restrictions will allow to expand the scope of possible applications of expert assessments. Prospects: the proposed approach and the set of techniques lay the methodological basis for creating a decision support system that will give specific recommendations to specialists involved in the development of program and methodology of the expertise at all stages of their activities. The introduction of such a system will greatly contribute to the simplification of various expertises, and will also increase confidence in their results.

1. Introduction

To date, most of the management and decision-making tasks have been successfully formalized at the process level, but not at the formal mathematical model level. Formal task sets and models are developed primarily for continuous control processes and systems. Management and decision-making in systems where the human factor, which is the cause of subjectivity and uncertainty, plays a significant role, is an area where so far there are more questions than answers. One of the established approaches to solving the problems of evaluation and prediction is expert assessment, which consists in analyzing the problem being solved, identifying its significant characteristics, and then constructing a mathematical or other model used later to assess the situation and develop a solution. In addition, in some cases, the use of expert assessment is much cheaper and more effective than other approaches [1]. Currently, promising methods and approaches have developed to a large extent, which can be used in the framework of expert assessment to build a formal model: analytic hierarchy process [2], methods of fuzzy set theory [3], the Bayesian approach [4], interval estimates [5].

It should be emphasized that experts themselves, who are specialists in specific problem areas, are not obliged to understand the specifics of the use of expert assessment methods. They should be considered as sources of information and «measuring instruments» on which the results of the expertise are based. This is all the more true, given the fact that to assess the quality and prospects of some developed information models that are planned to be used in the future, it is necessary to involve expert users who have not yet reached the age of majority.

In addition to the expert group, an important component of the reliability and adequacy of the results of the expertise is the efficiency of its preparation and provision, which is the responsibility of the working group. The activities of the working group can be compared with the activities of the designer, according to whose project the expertise itself will be carried out directly. However, the working group does not always have sufficiently qualified experts in the field of expert assessment, so it seems relevant to develop an information system of intellectual support for the activity of the working group in the process of preparing an expertise.

2. Organization of expertise

The expert commission that prepares and conducts the expertise consists of a working group and an expert group. To obtain reliable and adequate results of the expertise, effective and thoughtful interaction of these groups, performing various functions during the expertise, is required.

The working group ensures the work of the expert commission and performs the following functions at various stages of the expertise:

During the preparation phase, the working group:

- determines the number of experts in the expert group, selects them and forms an expert group;
- clarifies the object and purpose of the expertise, selects methods (analytical, modeling, survey, etc.), and also determines the need and possibility of conducting tests;
- specifies the methods and procedures for the assessment, determines the list of operations performed by experts and prepares the necessary program and methodology for the expertise.

At the stage of implementation of the expertise, the working group participates in solving the following tasks:

- adjustment (if necessary) of the expertise program and methodology;
- obtaining values of the estimated indicators;
- support of procedures for expert assessment of individual indicators.

At the final stage, the working group provides:

- obtaining integrated indicators that are the basis for results analysis and decision-making;
- analysis of the obtained integrated indicators and development of recommendations;
- documenting expertise results.

As can be seen, the working group, already at the stage of preparation, determines almost completely not only the process of conducting, but also to a large extent the quality of future expertise and, therefore, the adequacy of the expert assessments and the reliability of the conclusions drawn from these assessments.

Let's consider in more detail the support of the working group in solving the following important tasks:

- forming of expert group;
- creation of hierarchy of indicators;

- features and nuances of obtaining integrated indicators.

3. Forming of expert group

The composition of the expert group largely depends on the following factors:

- C – complexity and scale of the problem;
- R – necessary and available resources;
- E – efficiency of experts;
- S – successful participation in previous expertises;
- T – time allowed for the expertise.

Thus, the main factors affecting the composition of the expert group can be presented in the form of the next tuple:

$$G = \langle C, R, E, S, T \rangle .$$

At the same time, the complexity and scale of the problem determine which specialists should be involved as experts, that is, this factor generates problem-oriented tasks.

The existing resource constraints can affect the level of efficiency (competence, objectivity, etc.) of the experts involved. In addition, resource constraints often affect the number of experts involved.

The effectiveness of experts consists of the characteristics that the expert should have, in particular competence, objectivity, communicativity and ability to formalize his knowledge.

To assess how successfully the expert participates in previous expertises, it is necessary to organize the collection, storage and processing of data on the conducted expertises and their results. The successful participation of the expert in previous expertises (within a reasonable time interval) implies a sufficient proximity of his judgments and assessments to the real state of affairs.

The time factor can often turn out to be critical, since a sharp reduction in the timing and organization of the expertise can lead to the fact that a number of methods cannot be used due to their long duration. It can also affect the composition and size of the expert group.

Of all these factors, particular attention should be paid to the effectiveness of experts, depending on the following characteristics:

- Q – competence of the expert in the problem area;
- M – motivation for the expert to participate in the expertise;
- O – objectivity of the expert;
- P – personal qualities of the expert.

All characteristics that determine the expert's efficiency can also be represented in the form of the next tuple:

$$E = \langle Q, M, O, P \rangle .$$

It should be noted that the above characteristics of the expert's efficiency can also be broken down into separate components:

- competence of the expert in a problem area depends on his education, theoretical knowledge, practical experience;
- motivation of the expert depends to a large extent on the method of attracting the expert and his interest in participating in the expertise (request of colleagues, participation of authorities, prestige of participation, monetary remuneration, coercion, etc.);
- objectivity of the expert depends on his individual characteristics, and may also be due to external reasons (lobbying);
- personal qualities of the expert, which include psychophysical characteristics, erudition, sociability, etc.

It should be noted that if sufficient research is available on factors such as competence, motivation and personal qualities, there is practically no research on objectivity due to individual characteristics. Thus, for the selection of effective experts, the working group needs to use some procedure for the classification or assessment of experts according to the degree of their manifestation of the above characteristics.

4. Assessment of objectivity and stability of experts

Let there be a group of experts $A=\{A_1, A_2, ..., A_m\}$, as well as a number of expert assessments $x_1, x_2, \ldots, x_m$ of the value of *X*, where x_i is the assessment of the value of *X* given by the *i*-th expert. In accordance with the basic axiom of the group expert assessment:

$$\bar{x} \underset{m\to\infty}{\to} x, \tag{1}$$

where $\bar{x} = f(x_1, x_2, \ldots, x_m)$ is the integrated assessment obtained by aggregating the assessments of *m* experts; *x* is the true value of *X*.

Suppose that the *i*-th expert can estimate the value of *X* with correction for a possible error:

$$\varepsilon_i = x_i - x, \tag{2}$$

where ε_i is a random variable representing the error in the assessments of the *i*-th expert, which has a normal distribution $\varepsilon_i \sim N(\theta_i, \sigma_i)$ and contains two components:

$$\varepsilon_i = \theta_i + \xi_i, \tag{3}$$

where θ_i is a systematic error due to the individual characteristics of the *i*-th expert, equal to the mathematical expectation of ε_i; ξ_i is a random error caused by the influence of random factors and is a centered random variable.

The assumption of the normal distribution of errors of expert assessments is made on the basis of the following prerequisites: firstly, it is customary to use the normal distribution to describe various types of errors; secondly, this hypothesis ensures the simplicity of further analysis. The method for analyzing the objectivity of expert assessments presented below can be easily modified for those cases when, in the analyst's opinion, the distribution law of the error of expert assessments is different from the normal one.

For a certain object of assessment, let us calculate the numerical values of the errors made by experts when assessing this object. In accordance with (1), replace in (2) the true value of the estimated parameter by the integrated assessment: $\varepsilon_i = x_i - \bar{x}$. Let's move on to the relative error, to exclude the dependence of the error on the scale:

$$\delta_i = \varepsilon_i / r = (x_i - \bar{x})/r, \tag{4}$$

where *r* is the length of the scale. Thus, the relative errors of the assessments of all experts for a given object of assessment can be calculated: $\delta_1, \delta_2, \ldots, \delta_m$.

Let δ_{ij} be the error of the *i*-th expert when evaluating the *j*-th object. Then, according to the results of evaluating *n* objects by the *i*-th expert, it is possible to find errors $\delta_{i1}, \delta_{i2}, \ldots, \delta_{in}$ representing values of the random variable Δ_i, that is, the errors in the assessments of the *i*-th expert when using the considered criterion. Then we can calculate the average error of the expert's assessments for this criterion, which we will consider a systematic component of the error

$$\bar{\delta}_i = \frac{1}{n}\sum_{j=1}^{n}\delta_{ij} \underset{n\to\infty}{\to} \theta_i, \tag{5}$$

and also the corrected standard deviation of the error

$$\sigma_i = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(\delta_{ij} - \bar{\delta}_i)^2}. \tag{6}$$

5. Classification of experts

For each expert, there is a certain value of the systematic component of the error of his assessments (5), as well as the standard deviation of this error (6). Expressions (5) and (6) are characteristics of the objectivity of experts, according to the meanings of which experts can be classified.

According to the systematic component of the error, experts can be classified as follows:

- $\overline{\delta}_i \in [-\alpha_1;\ \alpha_1]$ – realist (*i*-th expert belongs to the class of realist experts);
- $\overline{\delta}_i \in [-\alpha_2;\ -\alpha_1)$ – moderate pessimist;
- $\overline{\delta}_i \in (\alpha_1;\ \alpha_2]$ – moderate optimist;
- $\overline{\delta}_i \in [-\alpha_3;\ -\alpha_2)$ – pessimist;
- $\overline{\delta}_i \in (\alpha_2;\ \alpha_3]$ – optimist;
- $\overline{\delta}_i \in [-\alpha_4;\ -\alpha_3)$ – excessive pessimist;
- $\overline{\delta}_i \in (\alpha_3;\ \alpha_4]$ – excessive optimist.

According to the random component of the error, experts can be classified as follows:

- $\sigma_i \in [0;\ \beta_1]$ – expert has a high stability;
- $\sigma_i \in (\beta_1;\beta_2]$ – expert has moderate stability;
- $\sigma_i \in (\beta_2;\beta_3]$ – expert has a low stability;
- $\sigma_i \in (\beta_3;+\infty)$ – expert is unstable.

Thus, for each assessment criterion, a classification of experts can be performed in accordance with the value of the systematic and random component of the error of their assessments by this criterion. The parameters α_j and ß$_j$ ($|\alpha_1|<|\alpha_2|<|\alpha_3|<|\alpha_4|$, $|\beta_1|<|\beta_2|<|\beta_3|$) are selected by the analyst based on the properties of the assessment scale and the requirements for experts in a particular expertise.

In some cases, the value of the systematic component of the assessment error does not allow to clearly classify one of the experts: $\overline{\delta}_{i*} \in (-\infty;\ -\alpha_4) \cup (\alpha_4;+\infty)$. It is suggested that such an expert be considered overly subjective and not take into account his assessments. However, in this case, the random component of the error of such an expert can be small, i.e. the expert is stable. Then, by analogy with [1], it is possible to calculate the correlation coefficient of this expert's assessments and integrated assessments obtained without taking into account the assessments of this expert. If this correlation coefficient turns out to be large enough, we can recalculate the assessments of this expert to eliminate his bias.

Obviously, if the random component of an expert's error allows this expert to be classified as an unstable expert, then his assessments should not be taken into account. A separate and rather difficult task is to identify affiliated experts whose bias is deliberate.

6. Creation of hierarchy of indicators

One of the tasks often performed by a working group when preparing an expertise is creation of hierarchy of indicators (assessment criteria), and these indicators can be related both to the object being assessed and, for example, to the experts themselves. Such a hierarchy can be built either by the working group itself or with the help of the expert group.

The most useful method of creation of hierarchy of indicators is implemented in two stages. At the first stage, a fully connected stepped hierarchy (initial) I_0 of interrelated indicators is built, which is presented to the specialists involved (figure 1(a)). Each of the experts can remove any links between elements of neighboring levels of the initial hierarchy and thereby form their own hierarchy, corresponding to his ideas about the relationship of indicators (figure 1(b)). At the second stage, the hierarchies received from experts are combined (figure 1(c)): $I_\Sigma = \bigcup_{k=1}^{m} I_k,\quad I_\Sigma \subset I_0$.

It should be noted that the union of hierarchies can be implemented not only in the form of a formal union of sets, but also in other ways. For example, if at least 2/3 of experts have voted for a given vertex (connection), then it will be included in the combined hierarchy. The choice of the method of merging hierarchies is the task of the working group, however, formal merging of hierarchies is the most preferable, since it covers all opinions and does not give grounds for subsequent adjustment of the hierarchy of indicators during the expertise.

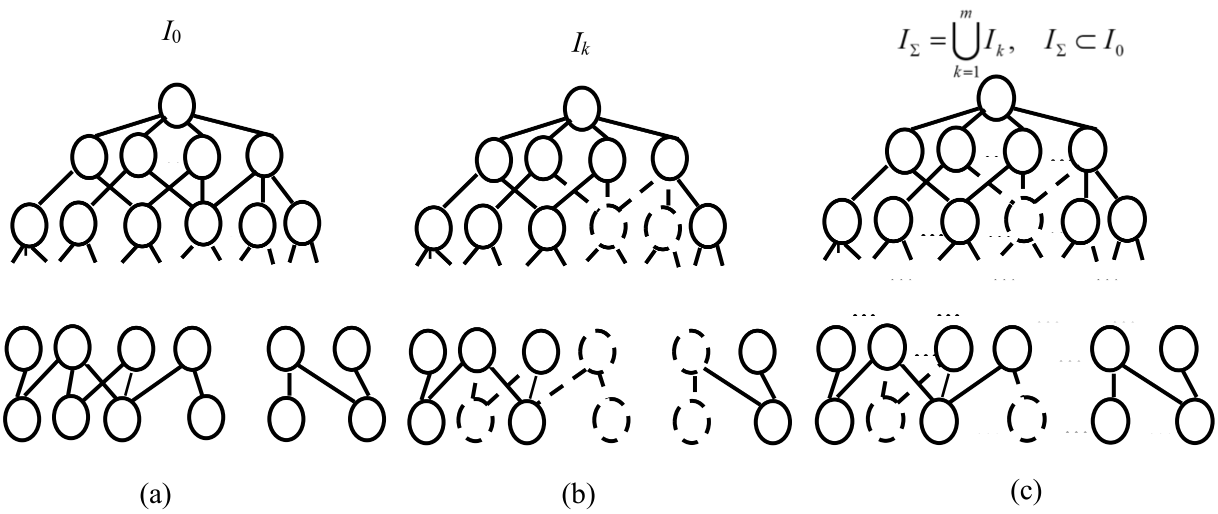

Figure 1. The initial fully connected hierarchy (a); hierarchy received from some expert (b); the resulting hierarchy, obtained after combining the opinions of all experts (c) (*the strokes mark the remote arcs and vertices in each of the particulars and in the resulting hierarchy*).

7. Features of obtaining integrated indicators

Obtaining integrated indicators is usually carried out using the weighted average method. The use of this method combines expert and calculation methods: on the basis of weight coefficients and values of assessments of single indicators the values of integrated indicators are determined by the calculation method, for example, according to one of the formulas given in table 1.

Table 1.Types and methods of calculating integrated indicators.

Type of assessment	Calculating method
Weighted average arithmetic	$\overline{z}_1 = f_1(x_1, x_2,..., x_m) = \sum_{i=1}^{m} q_i x_i \Big/ \sum_{i=1}^{m} q_i$
Weighted average geometric	$\overline{z}_2 = f_2(x_1, x_2,..., x_m) = (\prod_{i=1}^{m} x_i^{q_i})^{1/\sum_{i=1}^{m} q_i}$
Weighted average quadratic	$\overline{z}_3 = f_3(x_1, x_2,..., x_m) = \sqrt{\sum_{i=1}^{m} q_i x_i^2 \Big/ \sum_{i=1}^{m} q_i}$
Extreme pessimistic assessment	$\overline{z}_4 = f_4(x_1, x_2,..., x_m) = \min x_i$
Extreme optimistic assessment	$\overline{z}_5 = f_5(x_1, x_2,..., x_m) = \max x_i$

Since all assessments of single indicators are positive, $f_4 \le f_2 \le f_1 \le f_3 \le f_5$ inequality is always fulfilled.

The following cases may be the features of the methods for processing expert assessments:

1)implementation of the transition from one (initial) assessment scale to another scale, which due to certain properties is more convenient for experts or analysts;

2) discarding several pessimistic and optimistic assessments of experts, followed by obtaining the integrated assessment based on the remaining individual assessments.

7.1. *Case 1.*
It is clear that if during the processing of assessments a transition to the scale S^* was made, then the integrated assessment will belong to the same scale, which is not always convenient. For example, it may be necessary to provide integrated assessment to the experts who formed the initial vector of assessments in order to provide feedback during the expertise. Probably, such integrated assessment will be incomprehensible to the experts, because they used the scale S, and not the scale S^*. Therefore, the problem arises of transforming the integrated assessment obtained in the scale S^* into the original scale S. The simplest solution is to use the inverse transformation, which was used in the transition from the scale S to the scale S^*. But the application of the inverse transformation does not always give a correct result. The correctness depends on the nature of the transformation and the type of integrated assessment that undergoes a reverse transformation. In this case, speaking about the correctness, we mean the preservation of the integrated assessment on going from the scale S^* back to the scale S, that is, the fulfillment of the equality $L^{-1}\left(\bar{z}_i(L(X))\right) = \bar{z}_i\left(X\right)$, where L is any linear transformation.

The most frequent transformations used in the transition from one scale to another are the following linear transformations:

- Transformation of scale: $L_1(X) = \alpha X = (\alpha x_1, \alpha x_2, ..., \alpha x_m)$.
- Shift transformation: $L_2(X) = X + \beta = (x_1 + \beta, x_2 + \beta, ..., x_m + \beta)$.
- General linear transformation: $L_3(X) = \alpha X + \beta = (\alpha x_1 + \beta, \alpha x_2 + \beta, ..., \alpha x_m + \beta)$.

It is obvious that the above transformations (except for the case when α = 0) have inverse transformations. As was shown in [6], the scale transformation and the inverse transformation preserve the values of all integrated assessments. Also in [7] it is shown that when using the shift transformation and inverse transformation:

- geometric mean shifts by a greater value than the shift parameter, i.e. the assessment obtained by the inverse transformation will be overestimated;
- quadratic mean shifts by a smaller value than the shift parameter, i.e. the assessment obtained by the inverse transformation will be underestimated.

The general linear transformation is a superposition of scale and shift transformations, so its use is also not correct when the integrated assessment is obtained as the geometric or quadratic mean.

7.2. *Case 2.*
Consider a processing technique that assumes that several of the highest (k) and several of the lowest (l) expert assessments will be excluded before calculating the integrated assessment.

Assuming that the minimum assessment (a) and the maximum assessment (b) allowed by the scale are positive in [8] interval estimates of the deviations of the adjusted integrated assessments from the base integrated assessments are given.

When obtaining the integrated assessment by the method of simple arithmetic mean ($q_i = 1/m$), the difference between the base and corrected integrated assessments is in the interval $-\frac{l}{m}(b-a) \le f_1 - f_1^* \le \frac{k}{m}(b-a)$, and if the number of discarded highest and lowest assessments is the same ($l = k$), then the inequality $\left| f_1 - f_1^* \right| \le \frac{k}{m}(b-a)$ will hold.

When obtaining the integrated assessment by the geometric mean method, the difference between the base and corrected integrated assessments is in the interval $\left(\frac{b}{a}\right)^{-\frac{l}{m}} \le f_2 / f_2^* \le \left(\frac{b}{a}\right)^{\frac{k}{m}}$, and when switching to logarithmic scales: $-\frac{l}{m}(\ln b - \ln a) \le \ln f_2 - \ln f_2^* \le \frac{k}{m}(\ln b - \ln a)$.

In the case when the number of discarded highest and lowest assessments is the same ($l = k$), this expression will have the form: $\left|\ln f_2 - \ln f_2^*\right| \le \frac{k}{m}(\ln b - \ln a)$.

When obtaining the integrated assessment by the mean square method, the following inequality will be valid for the difference between the base and corrected integrated assessments:

$$-l \Big/ \left(\sqrt{\frac{la^2 + (m-l)b^2}{m}} + b\right) \le \frac{m\left(f_3 - f_3^*\right)}{b^2 - a^2} \le k \Big/ \left(\sqrt{\frac{(m-k)a^2 + kb^2}{m}} + a\right).$$

Thus, when developing the expertise methodology, the working group should take into account the following facts:

- when calculating the integrated assessment in the form of the arithmetic mean, the magnitude of the impact on it of the marginal assessment of one expert does not exceed 100 / m percent of the scale;
- integrated assessment in the form of the arithmetic mean is equally sensitive to overvalued and undervalued assessments of experts;
- when calculating the integrated assessment in the form of the geometric mean, the magnitude of the impact on it of the marginal assessment of one expert does not exceed 100 / m percent of the scale;
- integrated assessment in the form of quadratic mean is significantly more sensitive to overvalued than undervalued expert assessments.

8. Conclusions and prospects

The developed approach to the formation of the expert group is the basis for creating an expanded methodology that determines the main points of assessment and selection of experts for future expertise. It is planned to develop automated procedures that support the assessment and selection of experts for various expertises using a specialized database.

The proposed method for constructing a hierarchy of indicators can be used both for the future assessment of the studied object (process, system), and for a more in-depth assessment of prospective experts. The two-stage procedure for constructing the hierarchy of indicators will reduce the likelihood of experts' objections to the structure and composition of the main indicators, and thus avoid the need to adjust this structure at the stage of direct expertise, which is usually limited in time.

Analysis of the features and limitations of various approaches to obtaining integrated indicators based on the processing of individual expert assessments allows the working group to form a well-founded methodology for conducting an expertise at the early stages of preparation. Familiarization of experts with the prepared methodology will reduce the likelihood of disagreements between experts and the working group on the processing of expert assessments, as well as the likelihood of adjusting the program and methodology, and thereby increase the efficiency of the expertise itself.

All of the above can become the foundation for the formation and development of organizations specializing in the preparation and conduct of various expertises. On the basis of the stated results a system of intellectual support for the preparation and conducting of various expertises is being developed [9], which is part of the information system for decision support. The development and implementation of such a system will make it possible to solve many problems associated with expert assessment, including through the use of various methods of taking into account the uncertainty factor.

References

[1] Kreinovich V Ya 2019 Global Independence, Possible Local Dependence: Towards More Realistic Error Estimates for Indirect Measurements *Proc. of the XXII Int. Conf. on Soft Computing and Measurements* Saint Petersburg Russia 4-8

[2] Saaty R W 1987 The analytic hierarchy process: What it is and how it is used *Mathematical Modelling* **9** 161-76

[3] Zadeh L A 1965 Fuzzy sets *Information and Control* **8** 338–53

[4] Prokopchina S V 2018 Methods of mathematical statistics and econometrics under uncertainty based on the regularizing Bayesian approach *Soft measurements and calculations* **7** 30-51

[5] Utkin L V 2010 A class of pessimistic and optimistic regression models with interval and fuzzy data *Proc. of the 9th Int. Conf. on Application of Fuzzy Systems and Soft Computing* Prague Czech Republic 27-38

[6] Dutova E D, Nasarenko N A and Paderno P I 2016 Analysis of the influence of transformation and integration technology of the expert evaluations on the result *Proc. of the XIX Int. Conf. on Soft Computing and Measurements* Saint Petersburg Russia 21-24

[7] Burkov E A, Nazarenko N A, Nasser S S and Paderno P I 2019 Analysis of Correctness of Linear Transformations of Expert Estimates *Proc. of the XXII Int. Conf. on Soft Computing and Measurements* Saint Petersburg Russia 29-33

[8] Burkov E A, Lavrov E A, Paderno P I, Pasko N B and Siryk O E 2020 Analysis of Impact of Marginal Expert Assessments on Integrated Expert Assessment *Proc. of the XXIII Int. Conf. on Soft Computing and Measurements* Saint Petersburg Russia 11-14 (in Russian)

[9] Burkov E A, Lyubkin P L and Paderno P I 2017 Intellectual systems – the future of expert assessment *Proc. of the XX Int. Conf. on Soft Computing and Measurements* Saint Petersburg Russia 34-36