UDC 519.711

THE GENERAL IMPLEMENTATION OF DELPHI METHOD PROCEDURE WITH APPLICATION ON FUZZY DATA

A.A. DZUGAEV

This paper aims to propose mathematical construction of iterative consensus reaching procedure for expert opinions, based on classic Delphi method concept. This procedure can be applied in mathematical support for qualitative analysis within various types of foresight studies that require expert examinations. The construction of procedure is supported with methodical recommendations and application example based on fuzzy data.

The problems of scenarios creation for evolution processes in complex systems require thorough estimation of various properties for functional elements of systems being investigated. Expert panels and examinations relied on expert knowledge are known to be the popular and reliable tools for such estimations. At the same time, the problems of reaching consensus between experts and formation of consenting expert opinions are still greatly important, while these opinions are used as background for construction of alternate scenarios and determination of conditions for their implementations.

Solutions of foresight problems for industrial and economical systems often demand estimations for such complicated characteristics as competitiveness of production, economical effectiveness of enterprise, availability and feasibility of innovations [1]. All mentioned estimations are performed in conditions of incompleteness and uncertainty of incoming information and are remarkable for presence of some cryptic parameters, which cannot be measured immediately [2]. This situation leads to growing actuality of expert estimations, based upon knowledge, experience and insight of a human being specialist in his domain.

As a rule, no expert can give exact estimation in conditions of uncertainty, so his opinion is subjective and is characterized by certain degree of assurance. Owing to this, one of the most important characteristics of expert estimation procedures is the need for iterative modification and refinement of expert opinions, taking account of the earlier responses and feedback. These iterative refinements, if properly organized, may also solve the problem of reaching consensus within a group of experts, using technique known as Delphi.

It is necessary to mention, that in spite of longstanding experience and numerous applications of Delphi method [3, 4], a common holistic approach to Delphi examination procedure and mathematical support still does not exist. Most of previously described examples of Delphi applications either miss the description of data representation and processing technique, or this description is given in narrow data domain of concrete application. However, a lot of statistical data and qualitative information can be obtained from Delphi surveys, and data processing procedure developed for one type of examination will probably not satisfy the demands of another without considerable modification.

Research objectives. The Delphi method implementation for foresight problems requires that its mathematical support should be developed on the level of abstraction enough high to fit the whole variety of possible Delphi examinations no matter what problems or data domains they correspond to. Another objective for creation of this common Delphi procedure is to ensure high flexibility of data analysis algorithms that will be reached through independent multidimensional processing of Delphi surveys. This approach seems to be especially effective for investigation of unique, new and lately unknown objects of research, which parameters and organization are to be found. At the same time, for recently known objects of research, regarding to which experts has gained some experience, multidimensional analysis will allow reconsidering available knowledge in a new fashion and, finally, enlarging it.

Let us now consider the formal target setting, specified in [2].

Target setting. Let us consider a set (group) $E = \{e_k \mid k = \overline{1,K}\}$ of experts e_k, $k = \overline{1,K}$ that were asked to answer a set (survey) $Q = \{q_j \mid j = \overline{1,J}\}$ of questions q_j, $j = \overline{1,J}$. Every expert $e_k \in E$ is giving the answer to survey question $q_j \in Q$ as some opinion $\tilde{q}_{kj}$, $k = \overline{1,K}$, $j = \overline{1,J}$. Let us emphasize, that any opinion $\tilde{q}_{kj}$ from subset $\tilde{Q}_j = \{\tilde{q}_{kj} \mid k = \overline{1,K}\}$ of expert answers to survey question q_j is composed by expert e_k independently from other experts. We will also not specify here the nature of opinion $\tilde{q}_{kj} \in \tilde{Q}_j$. Requirements to opinion subsets $\tilde{Q}_j$, $j = \overline{1,J}$, namely existence of metric and quality functional on them, perhaps own for any $j = \overline{1,J}$, will be stated later.

It is necessary after some iteration (rounds) of examination and analysis of expert opinions (Fig. 1) to meet the specified consensus criterion and find for any survey question $q_j \in Q$ a group of experts, whose opinions on this question are forming consenting clusters. In every cluster, then, a resulting opinion should be selected at the end of every round that will represent an agreed decision for corresponding group of experts. Another important, yet not formal goal of examination is to retrieve comments from experts, containing reasoning and argumentations for their opinions and perform the subsequent anonymous exchange with this information, which will allow experts to refresh their knowledge and modify their opinions.

The final agreed opinions of the last round will represent the alternate answers of expert groups to each question of survey and may be used as background material for other methods of qualitative analysis [5], such as the Analytic Hierarchy Process, scenario creation and others.

Method of solution. Let us construct the analysis procedure for one round of examination. This procedure aims to analyze opinions $\tilde{q}_{kj}$ from all $k = \overline{1,K}$ experts to all $j = \overline{1,J}$ survey questions on each examination round. As we have mentioned above, all questions are formulated and answered independently, therefore expert opinions on each question will also be processed independently. Of course, the idea of independent processing sets up some claims on the structure of questions [6]. First, all questions have to be independent and there must not be any ambiguity. Second, there must not be any conditional statements, which make the primary question dependent on the fulfillment of a series of conditions. Questions where this occurs should be split into two or more separate questions.

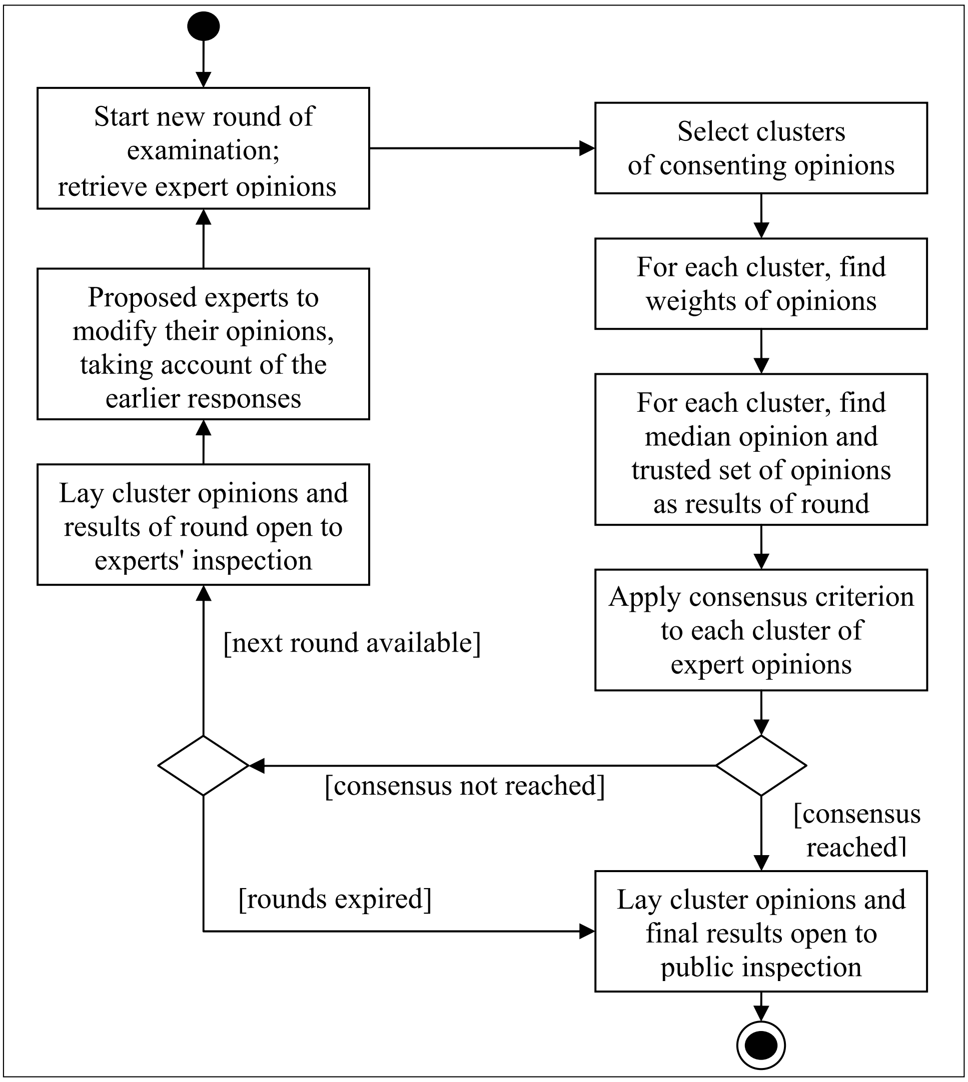

Fig. 1. The Delphi procedure flow diagram

Prior to opinion processing, we have to create all necessary tools of analysis. As will be shown below, there are just two things we need: the metric and the quality functional.

Introducing metric. To find out, how much expert opinions differ from each other, it is necessary to introduce metric on every opinion subset $\tilde{Q}_j = \{\tilde{q}_{kj} \mid k = \overline{1,K}\}$, and we are able to use different metrics for different $j = \overline{1,J}$. The metric $\rho_j : \tilde{Q}_j \times \tilde{Q}_j \to \Re$ allows for any pair of opinions $\tilde{q}_{ij}, \tilde{q}_{kj} \in \tilde{Q}_j$, $i, j = \overline{1,K}$ to define the measure of their distinction as distance

$r_{jik} = \rho_j(\tilde{q}_{ij}, \tilde{q}_{kj})$. The choice of metric is a crucial question that has to take into account the properties of data that represents expert opinions. In such a way we shift the features of data domain from method implementation to implementation of metric.

Introducing quality functional. Assigning weights to opinions aims to ensure their iterative refinement therefore requires mechanism of opinion quality determination. Opinions of highest quality should be considered more thoroughly, while the less quality opinions should be excluded from consideration, or at least their impact on analysis result should be reduced. We can create this mechanism by introducing the quality functional on every opinion subset $\tilde{Q}_j = \left\{\tilde{q}_{kj} \mid k = \overline{1,K}\right\}$, and again, we are able to use different functional for different $j = \overline{1,J}$. The functional $\omega_j : \tilde{Q}_j \to [0;1]$ allows defining the weight of any expert opinion $\tilde{q}_{kj} \in \tilde{Q}_j$, $j = \overline{1,K}$ as $w_{jk} = \omega_j(\tilde{q}_{kj})$, the weight of opinion is as high as its quality is considered to be. As you can see, we have again shifted the influence of data domain from method implementation to implementation of quality functional. The choice of quality functional has a strong impact on examination results, therefore it is better to rely on analytical characteristics of expert opinions, rather than on self-estimation of quality by experts. These self-estimations are mainly subjective and depend on personalities, their modesty and self-confidence.

Now, as we have all necessary tools, let us describe the method. Analysis procedure consists of the following five steps:

1. **Opinions clustering**. The aim of clustering is to prepare background for success of further analysis by means of dividing every opinion subset $\tilde{Q}_j = \left\{\tilde{q}_{kj} \mid k = \overline{1,K}\right\}$ into groups (clusters) of consenting opinions $C_{lj} \subseteq \tilde{Q}_j$, $l = \overline{1,L_j}$, $L_j \geq 1$. Clusters may be composed in various ways using metric ρ_j or, perhaps, some other properties of expert opinions. It is suggested, however, that in every cluster opinions should be enough close to each other, so that experts will be able to reach consensus among further iterations and, finally, agree upon one opinion. There probably will be some outlying opinions, and thorough argumentation will be required from their authors to share it with other experts. Every cluster will be analyzed with the object of median determination.

2. **Median determination**. Consider cluster C_{lj} containing N expert opinions $\tilde{q}_{ij}$, $\forall i = \overline{1,N}$ $\tilde{q}_{ij} \in \tilde{Q}_j$.

Median M_{lj} is defined as expert opinion from cluster C_{lj} least distant from the other opinions of C_{lj} in metric ρ_j (1).

$$M_{lj} = \tilde{q}^*_{pj} = \arg\min_{p=\overline{1,N}} \left(\sum_{i=1}^{N} \rho_j(\tilde{q}_{ij}, \tilde{q}_{pj}) \right). \tag{1}$$

Having built a symmetric matrix of distances $\mathbf{D}_{lj} = \{r_{jip}\}$ $i, p = \overline{1,N}$, where $r_{jip} = \rho_j(\tilde{q}_{ij}, \tilde{q}_{pj})$ for opinions $\tilde{q}_{ij} \in C_{lj}$ we are able to find median as opinion with minimal row sum.

For each examination round median represents the result of examination for group of experts that formed cluster C_{lj} on survey question q_j. The fact that examination result is represented by median, instead of average opinion that may be calculated in some possible ways, is of fundamental importance. First of all, this ensures commonality of method, since median determination requires nothing but metric, while calculation of average opinion will probably require some supplementary tools, probably dependent on the nature of data that represents opinions. Secondly, in general case there is never enough information for any unbiased averaging calculations on expert opinions, while median is always a reasoned opinion of some concrete expert.

3. **Weights assignment**. Before we build trusted set of opinions on cluster C_{lj} we should take into consideration quality of expert opinions. It is desirable to include in trusted set opinions with high quality, that is, with high weights, rather than with low weights and quality, which impact on examination result we are determined to reduce. Using the quality functional ω_j for opinions $\tilde{q}_{ij} \in C_{lj}$ we can calculate the weight vector $\mathbf{W}_{lj} = \{w_{ij}\}$ $i = \overline{1,N}$ $w_{ij} = \omega_j(\tilde{q}_{ij})$. As long as trusted set is built around the median, we may add to distance $r_{jiM} = \rho_j(\tilde{q}_{ji}, M_{lj})$ from opinion $\tilde{q}_{ji}$ to median M_{lj} a weight-determined correction factor, in order to increase distance from low quality opinions to median that will decrease their chances to get into trusted set. For example, to double the distance from median to opinion with weight 0 (the lowest quality) and to keep unchanged the distance to opinion with weight 1 (highest quality), we need a correction factor equal $(2 - w_{ij})$, $i = \overline{1,N}$, $j = \overline{1,J}$.

4. **Building trusted set**. The trusted set T_{lj} has to contain some fixed part of opinions from cluster C_{lj} least distant from its median M_{lj}. Opinions from trusted set are considered «consenting», while the rest opinions are «extreme». On the first round trusted set is built using a priori consent index S, that sets the ratio of consenting opinions to total amount of opinions in cluster. Thus, when $\mathrm{card}(T_{lj}) \le SN$ ($\mathrm{card}(T)$ means cardinality of the set, i.e. number of items in T), opinion $\tilde{q}_{ij} \in C_{lj}$, $i = \overline{1,N}$ is added to trusted set T_{lj}, if

$$\tilde{q}_{ij} = \underset{\tilde{q} \in C/T}{\arg\min} \left(\rho_j(\tilde{q}_{ij}, M_{lj})(2 - w_{ij})\right). \tag{2}$$

where $(2 - w_{ij})$ is weight-determined correction factor considered above. After trusted set is built, we define it's radius R_{lj} as distance from median M_{lj} to the most distant opinion in trusted set, considering correction factor $(2 - w_{ij})$.

$$R_{lj} = \max_{\tilde{q} \in T} \left(\rho_j(\tilde{q}_{ij}, M_{lj})(2 - w_{ij})\right). \tag{3}$$

On the next rounds of examination R_{lj} is fixed to its value R^1_{lj} on the first round and the principle is another: opinion $\tilde{q}_{ij} \in C_{lj}$, is added to trusted set, if

$$\rho_j\left(\widetilde{q}_{ij}, M_{lj}\right)\left(2 - w_{ij}\right) \le R^1_{lj} \qquad (4)$$

relatively the new found median.

5. **Applying consensus criterion**. On the first round of examination the trusted set radius R^1_{lj} is used for consensus criterion within each cluster C_{lj}. On the next rounds, consensus criterion is consent index S. Let us regard expert opinions in cluster C_{lj} as convergent, if cluster remains stable from round to round in the course of examination and for any round n holds $S_{n-1} \le S_n$. If the ratio S is decreasing, then expert opinions within given cluster are divergent and no consensus can be reached. If it is possible to pass another round, then divergent cluster will probably be divided between some other clusters. A signal to stop examination for cluster C_{lj} is when consent index S exceeds some prescribed threshold S^*, or it is just impossible to start another round for organizational reasons. With all this going on, the median of last round is taken as final agreed answer to survey question considered.

On the every new round of examination experts are acquainted with results of the previous round, including the fact of hitting or not hitting the trusted set with their opinions. Those experts, whose opinions are «extreme», that is clustering outliers or just out of trusted set, are asked to modify their opinions considering the judgment of the rest experts. Authors insisting on extreme opinions must share with others their arguments and reasoning, as feedback from them may be very important. Those who hit the trusted set may also reconsider their opinions under the influence of their colleagues. After new opinions are retrieved, median and trusted set are determined anew for all clusters on all survey questions. Results of examination and feedback are passed to experts anonymously, without concrete identification of authors. This aims to exclude obtrusion of opinions and influence of some authoritative personalities and may ensure far better convergence of opinions than any artificial methods and models.

Apparently, the global convergence of opinions never aims to be the primary outcome of Delphi examination. Variety of alternate opinions, estimations and views is highly appreciated in foresight and futures studies. Not only the opinions, but often reasoning, feedback and arguments given within experts' communication are important deliverables of the process. Another outcome of Delphi is the multiplication of expert knowledge, that is reached through experts cooperation and networking, when sharing views and feedback helps all stakeholders to learn more about the subject of studies.

Before we go on to application example for technique developed above, let us again accentuate, that mathematical methods and models, whatever sophisticated they could be, can not separately ensure the successful implementation of Delphi method. That is why especially the psychological process in connection with anonymous communication, together with reach feedback and argumentation has to be stressed.

Case study. Application on fuzzy data. Consider the application of Delphi method to expert opinions represented with fuzzy estimations. Let six experts ($K = 6$) estimate the risk of failure for some innovation activity. The segmented fuzzy estimation is built by expert through single-valued mapping the segments of risk value scale on segments of risk level scale (see Table), the scale segments having both quantitative and qualitative description.

Risk estimation for innovation activity

	risk value								
	qualitative description	lowest	lower	low	medium	high	higher	highest	
risk level	highest								[0,90; 1,00]
	higher		checked						[0,75; 0,90]
	high	checked		checked	checked				[0,60; 0,75]
	medium								[0,40; 0,60]
	low					checked	checked		[0,25; 0,40]
	lower							checked	[0,10; 0,25]
	lowest								[0,00; 0,10]
		[0,0; 0,1]	[0,1; 0,25]	[0,25; 0,4]	[0,4; 0,6]	[0,6; 0,75]	[0,75; 0,9]	[0,9; 1,0]	quantitative description

Consequently, the opinion $\tilde{q}_k$ of expert e_k on the risk q of given innovation activity is represented by fuzzy value μ_k, built by means of interpolation on discrete set of points chosen inside the segments checked by expert (see Figs 2–7). The fuzzy values are regarded here as a class of continuous functions $\mu : [0;1] \to [0;1]$, $\mu(x) \in C[0;1]$, that allows to introduce on them a standard L_1 metric $\rho(\mu_1, \mu_2) = \int_0^1 |\mu_1(x) - \mu_2(x)|\, dx$.

The quality functional for these fuzzy values will be stated as $\omega(\mu) = 1 - \rho(\mu, \breve{\mu})$, where $\breve{\mu}$ is a model for some kind of «perfect» estimation, representing Gaussian density function normalized to value area of fuzzy function $\mu(x)$ [7]. The model function mean (5) and dispersion (5a) are defined as follows:

$$a = \arg\max_{x \in [0,1]} \mu(x) \tag{5}$$

$$\sigma = \frac{1}{3}\left| a - \arg\min_{x \in [0,1]} \mu(x) \right| \tag{5a}$$

In such a case, the quality of opinion represented as fuzzy estimation is as high, as less fuzzy value function is distant from its model function in L_1 metric.

Consider now the sample expert opinions shown on Figs. 2–7. As there are just six opinions, they all will be included in a single cluster C, in which we have to choose median M and trusted set T.

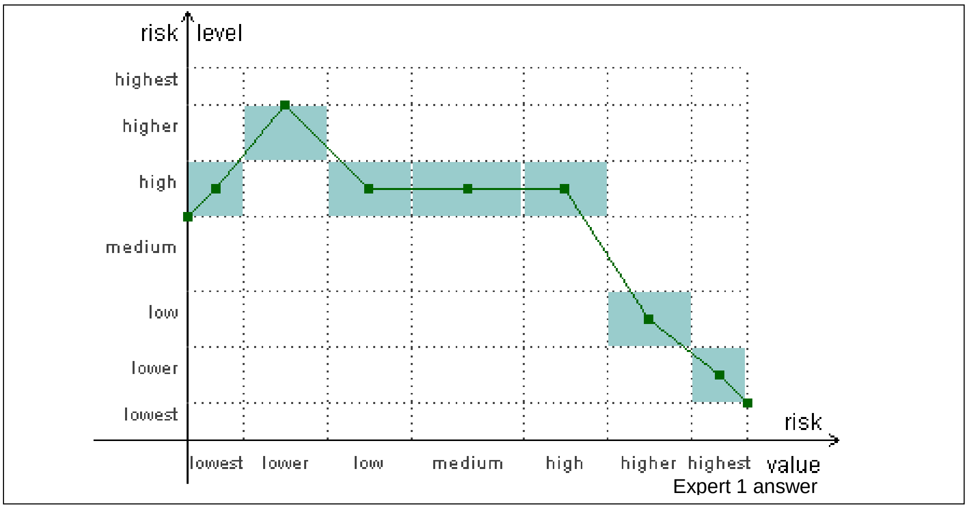

Fig. 2. Opinion of the first expert (trusted)

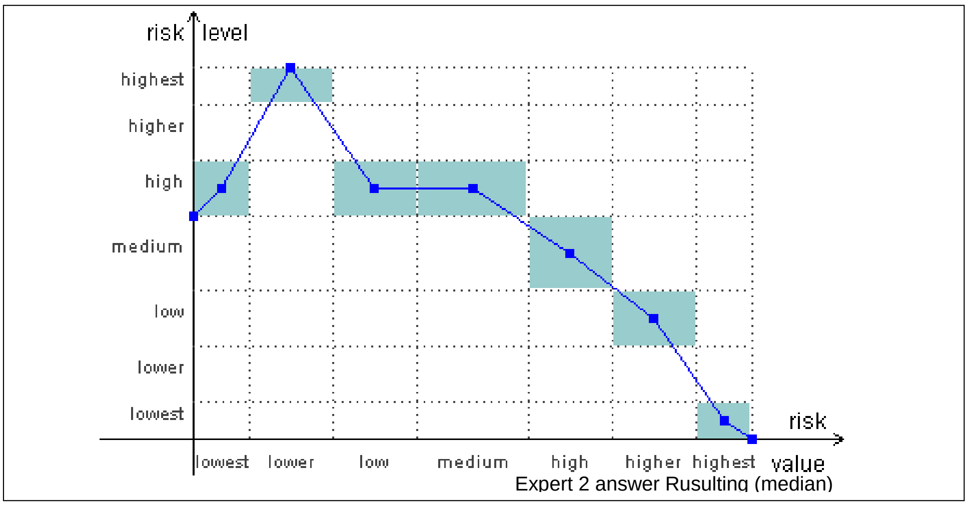

Fig. 3. Opinion of the second expert (median)

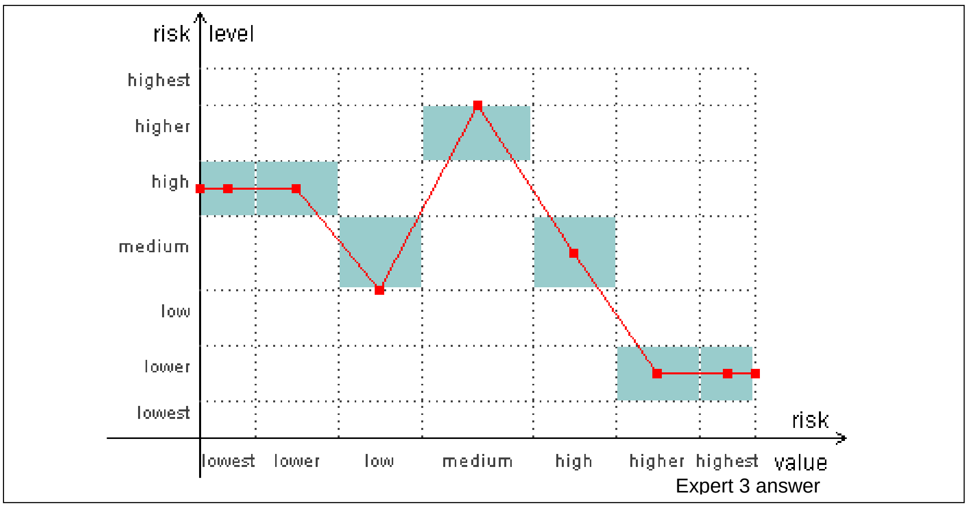

Fig. 4. Opinion of the third expert (not trusted)

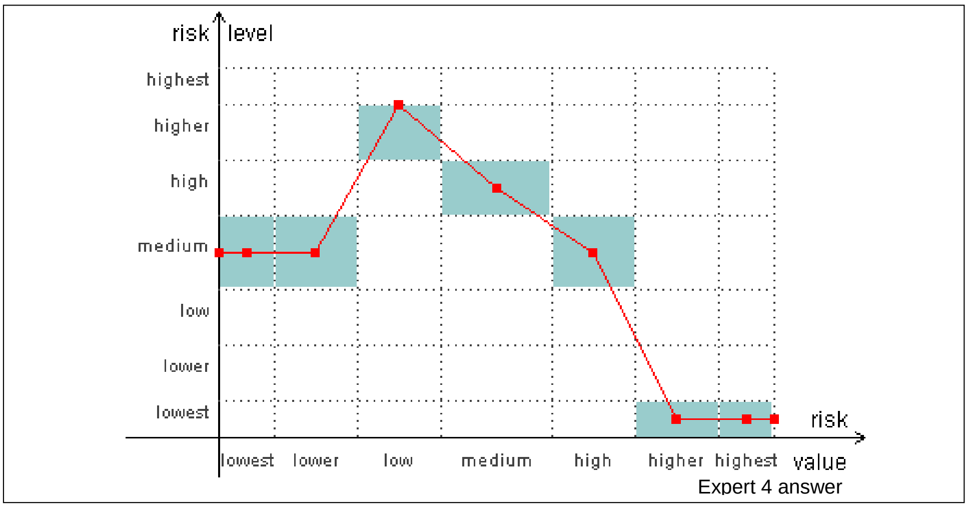

Fig. 5. Opinion of the fourth expert (not trusted)

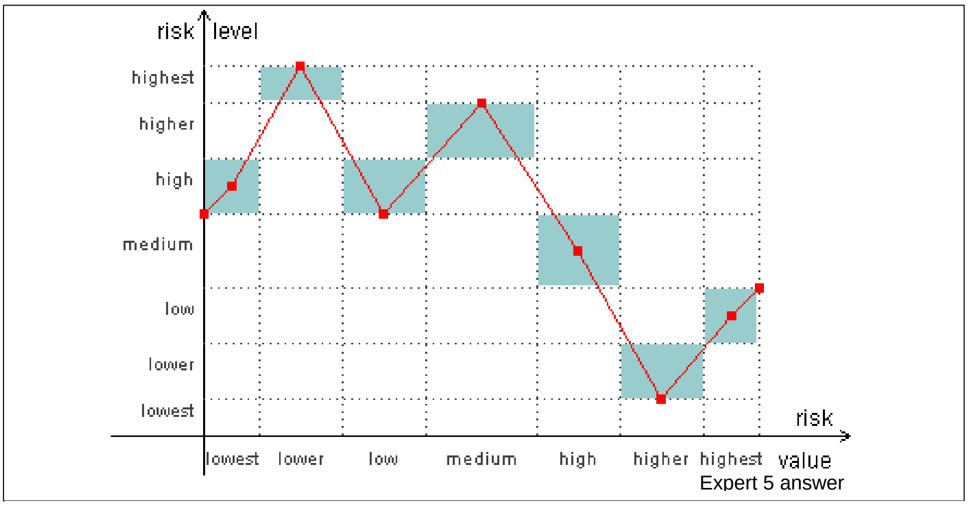

Fig. 6. Opinion of the fifth expert (not trusted)

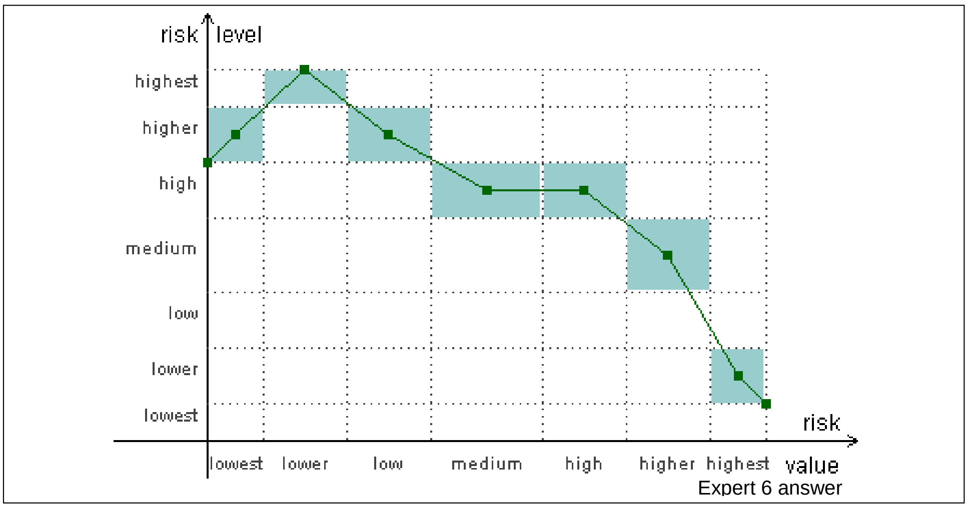

Fig. 7. Opinion of the sixth expert (trusted)

With the help of L_1 metric, the symmetric distance matrix $\mathbf{D} = \{\rho(\mu_i, \mu_k)\}$ $i, k = \overline{1, K}$ for cluster C is calculated as

0,0000	0,0556	0,1678	0,1878	0,1453	0,0791
0,0556	0,0000	0,1666	0,1622	0,1166	0,1072
0,1678	0,1666	0,0000	0,1956	0,1081	0,2431
0,1878	0,1622	0,1956	0,0000	0,2144	0,2181
0,1453	0,1166	0,1081	0,2144	0,0000	0,1969
0,0791	0,1072	0,2431	0,2181	0,1969	0,0000

The vector of row sums for matrix $\mathbf{D}$ equals

0,6356	0,6081	0,8813	0,9781	0,7813	0,8444

The median M by definition must have the least row sum, therefore the median in cluster C is the opinion of second expert, $M = \tilde{q}_2$.

The vector $\mathbf{D}^M = \{\rho(\tilde{q}_k, M)\}$, $k = \overline{1, K}$ of distances between cluster C opinions and median in terms of L_1 metric equals

0,0556	0,0000	0,1666	0,1622	0,1166	0,1072

As the result of weights assignment with the above defined quality functional ω we get the weights vector $\mathbf{W} = \{\omega(\mu_k)\}$, $k = \overline{1, K}$, equal

0,7688	0,8239	0,7150	0,8392	0,7305	0,7998

Adding the weight-determined correction factor $(2 - \omega(\mu_k))$, $k = \overline{1, K}$ into distances vector $\mathbf{D}^M$ we are transforming it to

0,0685	0,0000	0,2140	0,1883	0,1480	0,1287

With these weighted distances the trusted set T is calculated as the half ($S = 0{,}5$) of opinions in cluster C less distant from median. In this case, together with median, the opinions of the first and sixth experts are hitting trusted set. Consequently, the resulting opinion within considered six experts is the opinion of second expert $M = \tilde{q}_2$ (Fig. 3), and trusted set contains opinions of first, second and sixth experts, $T = \{\tilde{q}_1, \tilde{q}_2, \tilde{q}_6\}$. The consensus criterion for the next round of examination will be the trusted set radius $R = 0{,}1287$.

Fig. 8 shows the optimistic and pessimistic risk estimations built on opinion of second expert as median (Fig. 3). As we can see, well, the innovation is unlikely to fail.

Conclusion. The general implementation of Delphi method procedure proposed here is based upon the classic principles of experts examination, that were generalized and comprehended based upon the modern requirements to apparatus of qualitative analysis. The mathematical method for analysis of expert opinions is implemented on the higher level of abstraction, than required by any type of experts' examination that allows applying Delphi technique on any metric space whatever type of information it represents. Thanks to this approach both the mathematical method and its software implementation may be successfully applied to solve various problems of qualitative analysis, foresight and other fields of application of expert systems.

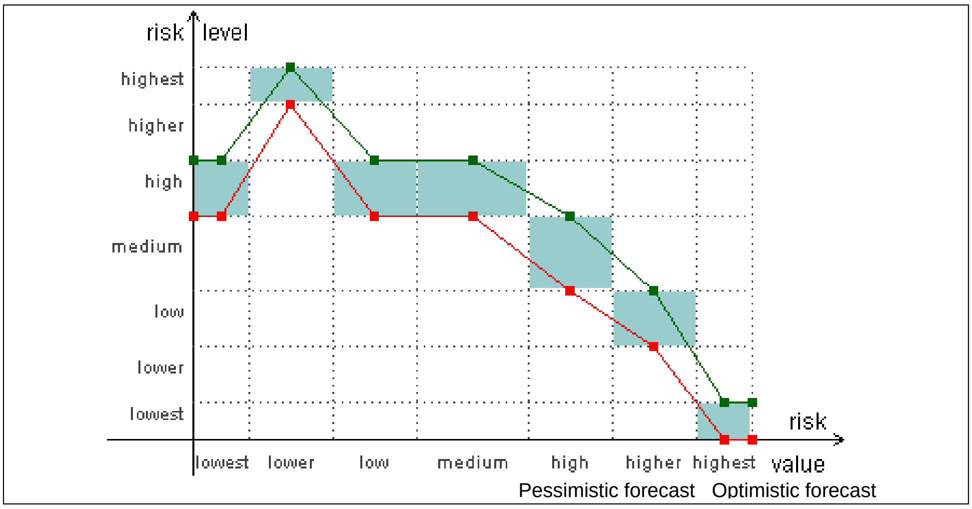

Fig. 8. The optimistic and pessimistic risk estimations

REFERENCES

1. *Zgurovsky M.Z.* The Scenario Analysis Platform as Methodological Base of the National Foresight Program of Ukraine. // System Research and Information Technologies. — 2003. — № 1. — P. 7–25.
2. *Pankratova N.D.* Software for Technology Foresight Problems with Reference to the Industry. // System Research and Information Technologies. — 2003. — № 1. — P. 26–33.
3. *Linstone H.A., Turoff M.* (eds.) The Delphi Method: Techniques and Applications. Addison-Wesley, Reading, Mass., 1975. — 570 p.
4. *Brockhaus W.L., Mickelsen J.F.* An analysis of prior Delphi applications and some observations on its future applicability. — Tech. Forecast and Soc. Change, 1977. — Ch.10. — P. 103–110.
5. *Loveridge D.* Foresight and Delphi Processes as Information Sources for Scenario Planning. Paper to the IIR Conference on «Scenario Planning» London, November, 1996. — 11 p. Available from http://www.les.man.ac.uk/PREST/People/Staff/ Denis_Loveridge.html
6. *Rauch W.* The Decision Delphi. — Tech. Forecast and Soc. Change, 1979. — Ch.15. — P. 157–170.
7. *Dzugaev A.O.* Fuzzy Estimations as Expert Opinions in Qualitative Analysis // Sixth International Conference on Applied Mathematics and Computer Science. National University of Lvov, Ukraine. — 2003. — P. 105–109.

Received 05.02.2004

From the Editorial Board: The article corresponds completely to submitted manuscript.