Multiplicity based algorithms for processing group multi-criteria expert assessments

S V Pronichkin$^{1,2,3}$, Y S Chernyshova$^2$, A V Kholstov$^4$ and I B Mamai$^2$

$^1$Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, 40, Vavilov Street, Moscow, 119333, Russia
$^2$Federal Center of Theoretical and Applied Sociology of Russian Academy of Sciences, 24/35, Krzhizhanovsky Street, Moscow, 117218, Russia
$^3$National University of Science and Technology "MISIS", 4, Leninsky Prospect, Moscow, 119049, Russia
$^4$N.N. Semenov Federal Research Center for Chemical Physics Russian Academy of Sciences, 4, Kosygina Street, Moscow, 119991, Russia

E-mail: pronichkin@mail.ru

Abstract. The paper discusses a new approach to the processing of verbal assessments, which express the opinions of experts on many criteria. The algorithm for verbal analysis of multi-feature objects expert assessments is proposed for identifying subgroups of experts with agreed opinions. It is proposed to use the multiplicity indices of multisets to calculate the coefficient of consistency. It is also proposed to use jointly the coefficients of consistency values to assess the consistency of the experts subgroups on one feature, and strive for the total maximization of the coefficients of consistency for all features. The complexity of the proposed algorithm is estimated. The results of calculations for the problem of choosing the best science-intensive technology are presented.

1. Introduction
There is a fairly wide range of problems in the study of complex systems and processes, where the objects under the study are characterized by many features, according to which verbal assessments are obtained from many experts. The main purpose of processing collective expert assessments is to obtain generalized data and identify new information contained in a latent form in expert assessments. The team of experts usually disagrees. In the absence of agreement among the experts, it is advisable to divide them into groups of similar opinions. In this regard, an important task to identify subgroups of experts with agreed opinions arises.

2. Mathematical formulation of the problem
The mathematical formulation of the problem is as follows.
Let $m$ experts evaluate $n$ objects according to $s$ criteria. The results of the assessment are presented in the form of values $x_{ijl}$, where $i$ is the number of the expert ($i=1,\ldots,m$), $j$ is the number of the object ($j=1,\ldots,n$), $l$ is the number of the criterion (feature) ($l=1,\ldots,s$). The relative importance of the criteria is known $w_l$, $l=1,\ldots,s$, $\sum_{l=1}^{s}w_l = 1$.

To obtain a breakdown of the original group of experts into subgroups with similar opinions for a
given criterion, it is customary to use the coefficients of consistency of opinions [1-3], which are calculated on the ranks, and allow assessing the consistency of opinions of experts in the group.

In [4, 5], the multiple qualitative consistency coefficient (1) is used, which is not rank-based.

\[
W_k = 1 - \frac{\sum_{j=1}^{J} \sum_{l_1=1}^{L} \sum_{l_2=1}^{L} |x_{l_1,j} - x_{l_2,j}|}{J \cdot L \cdot (L-1) \cdot (K-1)},
\]

where \( x_{l_1,j} \), \( x_{l_2,j} \) are the value \( j \)-th criterion of the object according to the results of the assessment by experts \( l_1, l_2 \);

\( L \) is the number of experts;

\( J \) is the number of criteria;

\( K \) is the basis of the features system quality.

The range of the coefficient (1) is the segment [0, 1]. A value of 0 corresponds to the maximum inconsistency, a value of 1 corresponds to a complete consistency.

The coefficient (1) is not free from shortcomings, since it does not take into account the distribution of expert assessments from one object to another and is based solely on numerical expert assessments according to the criteria obtained in the scale of intervals.

The numerical measurement of qualitative indicators, despite the seeming simplicity and obviousness, is not suitable for working with the qualitative characteristics that describe objects, since it contains a number of methodological defects.

It is practically impossible to assign a priori quantitative scales of assessments by comparing any numbers to qualitative factors so that they "correctly" express poorly formalized properties of objects and are equally understood by different people. Moreover, numerical estimates are inapplicable for measuring out-of-order indicators. Moreover, it was rigorously shown that the use of ordinal scales with different numerical rating scales, even according to the same criteria, can lead to completely different orderings and divisions into classes of the initial set of objects. At the same time, there is no substantive reasoning in favor of choosing a particular scale. Thus, only by arbitrarily changing the grades on the scales, it is possible to obtain any ranking and any classification of multi-feature objects. Finally, when using numerical rating scales, a generalized final indicator is usually used, which aggregates individual ratings (usually in the form of a sum, a weighted sum, or some averaged rating), according to the value of which the objects are compared. In this case, there is a mixing of dissimilar indicators, important and not important factors, assessments of different experts, which does not allow to single out the most significant indicators for selection.

Also, in [4, 5], the mechanism for assessing the significance of the proposed coefficient of multiple qualitative consistency is not proposed, i.e. strictly substantiating the statistical reliability of the resulting partition. It is important to note that no mechanism was proposed for identifying subgroups of experts with agreed opinions if the coefficient of multiple qualitative agreement is close to zero. At the same time, in [6-8], the algorithm is proposed in which subgroups with agreed opinions are sequentially excluded from the original group, and thus the original group is divided into subgroups with agreed opinions for a given criterion. This approach is not free from disadvantages.

The search for a subgroup of experts with agreed opinions stops as soon as the maximum value of the coefficient of agreement for the current number of experts in the subgroup turns out to be significant. This leads to the fact that the subgroup of experts with agreed opinions obtained first will be larger than the other subgroups, and in the general case, the division of the original group of experts into subgroups with agreed opinions will not be optimal in the sense of jointly maximizing the coefficients of agreement for subgroups of experts with agreed opinions for given criterion.

Only the value of the consistency coefficient for the subgroups of experts is taken into account, the range of the coefficient values variation, the dispersion of its values for a given criterion is not taken into account, which leads to unsatisfactory solutions, since the consistency of one subgroup of experts can be very different from the consistency of another.
The division of the original group of experts into subgroups with agreed opinions is carried out for a given (one) criterion. In the general case, it turns out that for each object, depending on the criterion, the set of experts rated it will be different, and the number of experts in the subgroup with similar opinions for one criterion may differ greatly from the number of experts in the subgroup with similar opinions for another criterion.

3. The proposed algorithm for processing expert estimates based on multiplicities

It is proposed to use multisets as a methodological basis for the algorithm for processing expert assessments [9, 10]. The theoretical model of a multiset is most suitable for structuring and analyzing a set of objects that are described by many verbal features, and may also be present in several versions. The assessment of an object is presented as a set of groups of elements $A_i = \{r_A(x) \cdot x | x \in X, r_A \in Z_r\}$, where $r_A : X \rightarrow Z_r = \{0,1,2,\ldots\}$ is the multiplicity function of the multiset, which determines the number of occurrences of an element $x_j \in X$ in the multiset $A_i$, according to the results of the $i$-th examination. Thus, it is proposed to use the multiplicity indices of multisets to calculate the coefficient of consistency. It is also proposed to use jointly the values of the coefficients of agreement for subgroups of experts, and strive for their total maximization at the same time, it is desirable that the agreement in the subgroups does not differ greatly. Each object is assessed by each expert according to many criteria, therefore, it is advisable to select subgroups of experts with agreed opinions on all the criteria, taking into account their importance.

It is proposed to construct an objective function $f(W_i^{(o)})$ to overcome the first and second disadvantages of existing algorithms. Which is the union (convolution) of the mean of the concordance coefficients and the standard deviation of the concordance coefficients for a given criterion $I$, where $i$ is the number of expert subgroups.

It is necessary to find such $W_i^{(o)}$, at which the coefficients of agreement mean value will be greater, and the value of the coefficients of agreement standard deviation will be less. The parameter values $\lambda_i \in [0,1]$ varying changes the properties of the objective function $f(W_i^{(o)})$ for $I$-th criterion, allowing, to a greater or lesser extent, depending on the value $\lambda_i$ to take into account the agreement coefficients mean values and the agreement coefficients standard deviation included in the objective function.

To overcome the third drawback, it is proposed to construct an objective function. Let us combine (carry out convolution) the mean and standard deviation of the objective function values $f(W_i^{(o)})$ taking into account the importance of the criteria. As the result, we get the following objective function:

$$ F(W^{(o)}) = (1-\lambda)\sum_{i=1}^{I} \frac{w_i f(W_i^{(o)})}{I} - \lambda \sqrt{\sum_{i=1}^{I} \frac{w_i^2 (\tilde{f}(W_i^{(o)}) - f(W_i^{(o)}))^2}{I}},$$ (2)

where $W^{(o)} = (W_1^{(o)},\ldots,W_I^{(o)})$ are vectors of the coefficients of agreement obtained for dividing the original group of experts into $I$ subgroups $E^{(o)} = (E_1,\ldots,E_{g_1},\ldots,E_{g_{i+1}},\ldots,E_{g_I})$;

$\tilde{f}(W_i^{(o)})$ is the weighted average value of functions $f(W_i^{(o)})$;

$w_i$ is the importance of the criteria, $i=1,\ldots,I$, $\sum_{i=1}^{I} w_i = 1$;

$\lambda$ is the parameter (weight coefficient).

The task of identifying subgroups of experts with agreed opinions is formalized in the form of the objective function (3):
where $W_{\text{sign}}^{(t)}$ is the set of significant coefficients of consistency, $W_{\text{sign}}^{(t)} = \{W_i(c_u) : W_i(c_u) \geq W_{ti}(c_u), u = 1, \ldots, t, i = 1, \ldots, I\}$; $W_{ti}(c_u)$ is the theoretical value of the coefficient of agreement for a subgroup of experts $u$.

It is necessary to find a value of $E_{\text{opt}}^{(t)}$, that corresponds to $W_{\text{opt}}^{(t)}$, from the equation (3).

A formal description of the algorithm for solving the equation (3) is:

1. Set $p_{\text{err}}(i)$ – probability of an error, $\lambda$, $i = 1, \ldots, I$.
2. $\lambda = 0$, $t_1 = 1$.
3. Complete matrices of multiplicities of multisets $\|c_p\|$, $d = 1, \ldots, m$, $p = 1, \ldots, P$, $i = 1, \ldots, I$ sequentially in different combinations to divide into $t_1$ groups of rows, rows in a group $c_u = 2, \ldots, m - 2(t_1 - 1)$, $u = 1, \ldots, t_1$, if $m - 2t_1 \geq 1$, then $\sum_{u=1}^{t_1} c_u = m - 1, m$, otherwise $\sum_{u=1}^{t_1} c_u = m$. For each set of rows, compute $W_{\text{sign}}^{(t_1)} = (W_{1}(c_1), \ldots, W_{t_1}(c_{t_1}))$, $i = 1, \ldots, I$ and compare $W_i(c_u)$ with $W_{ti}(c_u)$, if $W_i(c_u) \geq W_{ti}(c_u)$, $u = 1, \ldots, t_1$, $i = 1, \ldots, I$, then compute $F(W_{\text{sign}}^{(t_1)})$. Find $F(W_{\text{sign}}^{(t_1)})$ with the maximum value of $F(W_{\text{sign}}^{(t_1)})$.
4. $t_1 = t_1 + 1$.
5. Check the inequation $t_1 \leq \left\lfloor \frac{m}{2} \right\rfloor$, if the inequation is satisfied go to step 3.
6. Find $F(W_{\text{opt}}^{(1)})$ with the maximum value of $F(W_{\text{opt}}^{(t)})$. Remember $\lambda$, $t$, $W_{\text{opt}}^{(t)} = (W_1^{(t)}c_1, \ldots, W_t^{(t)}c_t)$ and the corresponding matrix of multiplicities of multisets $\|c_p\|$, $d = 1, \ldots, m$, $P = 1, \ldots, P$, $i = 1, \ldots, I$ with the agreed opinions of experts.
7. $\lambda = \lambda + 0.1$.
8. Check the inequation $\lambda \leq 1$, if the inequation is satisfied go to step 3.
9. If the solution is not found go to step 1.

In the proposed algorithm, the values of the parameters $\lambda$, $l = 1, \ldots, s$ can be set immediately before the calculation $F(W_{\text{sign}}^{(t)})$ (see step 3). As an example, the step of changing the parameter $\lambda$ is 0.1. The size of the step of changing the parameter $\lambda$, $h_\lambda$ is set by the decision maker (DM), then the DM examines the obtained solutions (3) at $\lambda = \{0; h_\lambda; \ldots; 1\}$ and chooses $\lambda$ those corresponding to his idea of the quality of the solution.

The complexity estimate of the above algorithm is exponential $O\left(\sum_{t_1=1}^{[\frac{m}{2}]} C_m^{c_1} \cdot C_{m-c_1}^{c_2} \cdot \ldots \cdot C_{m-c_1-c_2-\ldots-c_{t_1-1}}^{c_{t_1}}\right)$, where $C_m^{c_i}$ is the number of combinations of $m$ different elements by $c_1$ elements, the order of the elements does not matter. Despite the fact that the complexity estimate is exponential, nevertheless, there is a rule of thumb for the number of experts in a group, it should vary from 2 to 12. Within the specified limits, the algorithm gives an exact solution without significant time costs. Specific calculation times within the specified limits are from 0.127 to 4.643 sec.
4. Calculation results for the problem of choosing the best technology

The modern stone processing industry uses various technologies for the processing of rocks, which can be divided into two large classes [11]: mechanical and physical and mechanical ones.

Currently, the most widespread technologies are based on the mechanical method of stone processing, such as chipping, cutting and impact destruction, which are still the most common.

Among the physical and technical methods of stone processing, the technologies of processing with thermosetting gas burners, high-frequency currents, high-speed water jet and plasma destruction of rocks are widely used.

![Figure 1](image.png)

**Figure 1.** Dependence of the partition (a group of experts with agreed opinions) on the values of the parameters

Nowadays, new technologies for processing rocks are being actively developed [12, 13]. The task of choosing the best science-intensive technology is multi-criteria and means that during its implementation all the necessary restrictions of a technical, environmental, social, financial and other nature are observed. The indicators for assessing the effectiveness of technology can be a set of indicators of a different nature. They can be expressed in both qualitative and quantitative form. However, only the use of qualitative indicators allows obtaining a comprehensive expert assessment. A qualitative approach to solving the problem of choosing a technology makes it possible to overcome the difficulty, which consists in the absence of methods for obtaining objective quantitative indicators and such indicators as: the contribution of the developed technology to the main direction of the enterprise; the likelihood of success in the implementation of the technology; increment of the technology readiness level for the planned period. This task can be solved by introducing an integral system for assessing the technologies readiness levels, including not only the assessment itself, but also the issues of system integration of technology. Thus, the advantage of qualitative indicators is the possibility of obtaining a broader assessment due to the fact that the assessment is formulated not in the form of one or several quantitative characteristics, but in the form of a qualitative description. In particular, the scale of the technologies readiness level has found a wide application in world practice.
[14, 15], which is a set of criteria for evaluating a technology from the point of view of its implementation, starting from an idea and ending with a ready-made prototype. The scale of the readiness level allows to consistently compare various technologies readiness level according to such criteria as: the degree of support for the technology development program; the effectiveness of theoretical and empirical design solutions; completeness of the technology basic element; opportunities for improving technology, etc. Moreover, at any of the technology readiness levels, collective expert assessments are used.

The proposed algorithm was used to solve the problem of choosing the best technology for decorative stone processing. The assessment was carried out with the involvement of seven experts. The calculation results depend on the choice of parameters (Figure 1).

The decrease in the value $\lambda_i$ leads to the increase in the number of groups and, accordingly, to their small composition. The value $\lambda_i$ increasing has the opposite effect, and one has to change the level of significance (increase) because the coefficients of consistency become insignificant for the selected subgroups.

By choosing $\lambda$ the decision maker gets the result corresponding to his idea of the quality of the solution. This visibility allows the decision maker to study the resulting set of solutions, to formulate their requirements more clearly by comparing different solutions.

5. Conclusion
The main purpose of processing collective expert assessments is to obtain generalized data and identify new information contained in a latent form in expert assessments. The study of the opinions of experts subgroups with agreed opinions on many criteria can help a decision-maker to formulate selection strategies that are more adequate to reality, make his decisions more grounded and reasonable. In the absence of agreement among the experts, it is advisable to divide them into groups of similar opinions. A new algorithm for processing verbal assessments is proposed, in which the experts’ opinions are expressed on many criteria. As the methodological basis, multisets are used that are most suitable for structuring and analyzing a set of objects that are described by many verbal features, and may also be present in several versions. The proposed algorithm was used to solve the problem of choosing the best science-intensive technology. The use of the proposed algorithm for processing expert assessments based on the metric spaces of multisets makes it possible to more adequately structure and analyze a set of technologies that are described by many verbal features and are present in several versions than in the case of using existing approaches.

6. Acknowledgments
The article is prepared with the financial support of the Russian Science Foundation, project № 19-78-10055.

References
[1] Grzegorzewski P 2006 Computational Statistics & Data Analysis 51 314-22
[2] Kuznetsov M and Strijov V 2014 Expert Systems with Applications 41 1988-96
[3] Croux C and Dehon C 2010 Statistical Methods & Applications 19 497-515
[4] Tong F, Tang S, Irby B and Guerrero C 2020 International Journal of Educational Research 99 101514
[5] Zhao Y 2016 Journal of University of Science and Technology of China 46 165-72
[6] Xu L, Wang Z and Dai Y 2010 Journal of Nanjing University of Science and Technology (Natural Science) 34 30-33
[7] Alvarez E, Fernandez P, Encomienda F and Munoz J 2015 Journal of King Saud University-Science 27 271-7
[8] Mollick E and Nanda R 2015 Management Science 62 1533-53
[9] Blizard W 1993 Theoretical Computer Science 110 79-98
[10] Petrovsky A 2019 Group Verbal Decision Analysis (Moscow: Nauka)
[11] Zhu H, Li X and Zhuang X 2011 *Journal of Rock Mechanics and Geotechnical Engineering* 3 220-33
[12] Lawal A and Kwon S 2020 *Journal of Rock Mechanics and Geotechnical Engineering* 13 248-66
[13] Wang Q, Gao S, Li S and Jiang Y 2018 *Tunnelling and Underground Space Technology* 73 287-94
[14] Martinez-Plumed F, Gomez E and Hernandez-Orallo J 2020 *Telematics and Informatics* 58 101525
[15] Javed K, Gouriveau R and Zerhouni N 2017 *Mechanical Systems and Signal Processing* 94 214-236