ISSN 2070-0482, Mathematical Models and Computer Simulations, 2020, Vol. 12, No. 6, pp. 907–914. © Pleiades Publishing, Ltd., 2020.
Russian Text © The Author(s), 2020, published in Matematicheskoe Modelirovanie, 2020, Vol. 32, No. 3, pp. 115–126.

Factor Modeling for Innovative Enterprises

B. N. Chetverushkin[a] and V. A. Sudakov[a, b, *]

[a] *Institute of Applied Mathematics, Russian Academy of Sciences, Moscow, Russia*

[b] *Plekhanov Russian University of Economics, Moscow, Russia*

**e-mail: sudakov@ws-dss.com*

Received July 15, 2019; revised July 15, 2019; accepted October 21, 2019

Abstract—The concept of a factor model based on frames is proposed. Factors are considered as slot values of the corresponding frames, which makes it possible to reduce the complexity of the developed model. The method for calculating the factor values by searching for the eigenvalues of the pairwise comparison matrices of the extent of the slots' influence is proposed. A procedure allowing us to see the spread of the influence of factors along all possible paths of the influence graph is developed. As a result of the calculations, the most influential factors on the efficiency of the entire modeled system are determined. It is shown that it is possible to use the factor model for estimating the activity of innovative enterprises and for selecting an electronic flight bag (EFB) for the crew. This model can be used to forecast the development of complex dynamic systems with feedback, as well as for evaluating and making management decisions on the innovative activity of enterprises under conditions of weakly structured information.

Keywords: factor model, frame, slot, eigenvector, eigenvalues, innovative enterprise

DOI: 10.1134/S2070048220060058

1. INTRODUCTION

Modeling the activity of innovative enterprises requires simultaneously accounting for a large number of heterogeneous indicators. Due to the weak structure of the problem and the lack of representative statistical data on innovative activity, it is difficult to construct a quantitative model.

One approach to modeling such systems is to build a factor model [1]. It is a directed graph in which the vertices are the factors-indicators that determine the state of the described system, while the arcs are the relations between different factors. An example of such a graph for the case of five factors $x_0, x_1, \ldots, x_4$ is shown in Fig. 1. The loops at some vertices show the influence of the factor on itself; i.e., there are factors whose value depends not only on the influencing factors but also on the factor's previous value.

The degree of influence of one factor on another is reflected differently. In the simplest case, it is characterized only by a plus or minus sign, which indicates that an increase in the value of one factor leads to

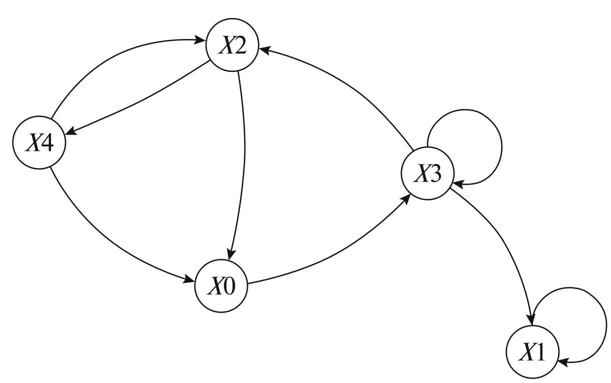

Fig. 1. Factor model graph.

an increase or decrease in the value of another factor. Such models are called cognitive models. They are usually based on expert judgments.

In more complicated variants, the coefficient of influence characterizing the degree of influence of one factor on another is assigned to each arc. The values of these coefficients can be obtained based on expert judgments or determined based on experience, including probabilistic estimates [2]. There may be more complicated cases when the coefficients depend on the factors themselves, i.e., nonlinear dependence. In this article, the problem with static coefficients of influence is mainly discussed.

2. DESCRIPTION OF THE METHOD

The factor model can be formalized in matrix form:

$$\mathbf{X} = A\mathbf{X} + \mathbf{F}. \tag{1}$$

Here, $\mathbf{X}$ is the vector of values of factors with dimension n, A is the matrix of the coefficients of the influence between corresponding factors, and $\mathbf{F}$ is the vector of values characterizing the external influence.

Factor values can be found by using an iterative procedure:

$$\mathbf{X}^{k+1} = A\mathbf{X}^{k} + \mathbf{F}, \tag{2}$$

where k is the iteration number.

The convergence rate of this procedure (and the possibility of convergence) is determined by the eigenvalues of matrix A.

In many cases, when the number of factors is large, we have to deal with a high dimension of the problem being solved. The volume of calculations in procedure (2) and the need for computational resources for it increase greatly. It becomes difficult to perceive the visual representation of the graph. We propose modeling factors based on the concept of frames to solve this situation.

A frame is a way of presenting knowledge in the artificial intellignece (AI) [3] and corresponds to some notion of the real world. A frame consists of slots which contain structured knowledge. The value of a slot is the value of a factor at a certain time. Correspondingly, the frame is a concept from the domain which aggregates the number of slots by a semantic feature. The process of transition from slots to the frame is similar to the condensation of the graph [4].

The frame structure makes it possible to reduce the complexity of the system's description. Without detailed consideration of the structure of the frame, we can study how the slots of one frame influence the slots of another frame. By analogy with graph condensation, the concentration of several vertex-factors in a single node of the frame makes it possible to obtain a graph of a smaller dimension. Internally, the influence of slots on each other within the frame is defined by the loops at the corresponding vertices.

Let us introduce the following designations: $i = 1, \ldots, m$ is the frame number, $I = \{i\}$ is the set of frame numbers, J_i is the tuple of numbers of the factors of the ith frame, q_i is the number of factors of the ith frame, (i_1, i_2) is the arc indicating the influence of the i_1th frame on the i_2th frame

Let us consider the directed graph of frame relations $G = (I, \{(i_1, i_2)\})$. The inversely symmetric matrix of pairwise comparisons of the degree of influence of $j \in J_{i1}$ on j^* is constructed in it for each arc (i_1, i_2) and each factor $j^* \in J_{i2}$:

$$U_{i_1 j^*} = \begin{Vmatrix} 1 & u_{12}^{j^*} & u_{13}^{j^*} & u_{14}^{j^*} & \ldots & u_{1q_{i_1}}^{j^*} \\ 1/u_{12}^{j^*} & 1 & u_{23}^{j^*} & u_{24}^{j^*} & \ldots & u_{2q_{i_1}}^{j^*} \\ 1/u_{13}^{j^*} & 1/u_{23}^{j^*} & 1 & u_{34}^{j^*} & \ldots & u_{3q_{i_1}}^{j^*} \\ 1/u_{14}^{jj^*} & 1/u_{24}^{j^*} & 1/u_{34}^{j^*} & 1 & \ldots & u_{4q_{i_1}}^{j^*} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1/u_{1q_{i_1}}^{j^*} & 1/u_{2q_{i_1}}^{j^*} & 1/u_{3q_{i_1}}^{j^*} & 1/u_{4q_{i_1}}^{j^*} & \ldots & 1 \end{Vmatrix}, \tag{3}$$

where $u_{j_1 j_2}^{j^*}$ shows how many times factor j_1 influences factor j^* more than factor j_2.

From the following equation,

$$(\lambda_{i_1 j^*} I - U_{i_1 j^*})\mathbf{w}_{i_1 j^*} = 0 \tag{4}$$

let us find the maximum eigenvalue $\lambda_{i_1 j^*}$ and the corresponding vector of factors $w_{i_1 j^*}$ of influence.

This method is similar to the methods of analytical hierarchy and network proposed by T. Saaty [5, 6]. In this case, not only the criteria characterizing the performance of the modeled system but also other characteristics of the considered system, e.g., resource limitations, can be factors. The notion of frames can be interpreted quite broadly; e.g., it is possible to introduce an inheritance hierarchy that makes it possible to apply the same types of slots to frames with a common ancestor.

Let us construct a generalized influence matrix A by all factors:

$$A = \begin{Vmatrix} w_{i_1 1}^{1} & 0 & \dots & \dots & 0 \\ w_{i_1 1}^{2} & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ w_{i_1 1}^{q_{i_1}} & 0 & \dots & \dots & 0 \\ 0 & w_{i_2 2}^{1} & \dots & \dots & 0 \\ 0 & w_{i_2 2}^{2} & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & w_{i_2 2}^{q_{i_2}} & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & w_{i_n n}^{1} \\ 0 & 0 & \dots & \dots & w_{i_n n}^{2} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & w_{i_n n}^{q_{i_n}} \end{Vmatrix}. \tag{5}$$

Let us denote the matrix elements $A = \|a_{ij}\|$; they are defined by the following formulas:

$$a_{ij} = \begin{cases} w_{i_1 j}^{i^*}, & \text{if } \exists i_1 \exists i_2 : ((i_1, i_2) \in G) \wedge (i \in J_{i_1}) \wedge (j \in J_{i_2}), \\ 0, & \text{otherwise}, \end{cases} \tag{6}$$

where i^* is the serial number of factor i in tuple J_{i1}.

The corresponding j columns contain vectors $\mathbf{w}_{i_k j}$ in which there is influence of frame i_k. The rest of the elements are filled with zeros.

Let us normalize the initial factor influence coefficients (elements of matrix A) so that the sum of the elements in the column of the new matrix A^* is 1:

$$\sum_{k=1}^{n} a_{kj}^{*} = 1, \quad \forall j. \tag{7}$$

Here, the normalized elements of the new matrix $A^* = \left\|a_{kj}^{*}\right\|$ are found using the following ratio:

$$A^* = \operatorname{norm}\langle A \rangle = A \begin{Vmatrix} \sum_{k=1}^{n} a_{k1} & 0 & \dots & 0 \\ 0 & \sum_{k=1}^{n} a_{k2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sum_{k=1}^{n} a_{kn} \end{Vmatrix}. \tag{8}$$

Let us construct a sequence of matrices defined by the following ratio:

$$A^1 = A^*, \quad A^h = \operatorname{norm}\langle A^{(h-1)} A^* \rangle \quad \text{for} \quad h = 2, 3, \dots. \tag{9}$$

This procedure makes it possible to see the spread of the influence of factors along all possible paths of the influence graph to which the matrix A corresponds. Its elements show only the direct influence of each system factor on all other factors. Frame slots can be other frames, which leads to the creation of a hierarchy of factors influencing each other. Factors can influence each other indirectly through some transition factors. All the possible influence paths through transition factors should be considered. To do this, let us construct a sequence of influence matrices: A^1, A^2, A^3, A^4, etc.

Consider the limit

$$\lim_{k\to\infty}\sum_{h=1}^{k} A^h/k. \tag{10}$$

If a certain sequence converges to the limit, its Cesàro sum converges to the same limit [7]. Since sequence (9) is given by the integer values of the matrix power h, it is sufficient to determine the limit value of A^h. It is possible that the sequence does not converge to a single limit but the average of Cesàro sums (10) corresponding to different limits of the sequence gives a single limit. Both of these cases may occur when raising matrices to a power. The Jordan canonical form for the stochastic matrix A helps to make sure that this limit exists. It is known that A is similar to its the Jordan matrix Y [8] if there is a nondegenerate matrix P such that

$$Y = PAP^{-1}.$$

Consequently, raising A to such powers h that A^h ceases to change with an accuracy up to the predetermined error is equivalent to raising Y to the same power h. Each square matrix is associated with a single Jordan matrix, which consists of square blocks, the main diagonals of which lie on the main diagonal of the lower triangular matrix A. All matrix elements outside these blocks are 0. Diagonal elements in each block are equal to their eigenvalue A, elements above the main diagonal are 1, and those below the main diagonal are 0. Matrix A is the direct sum of the blocks of the Jordan matrix. We can argue that the limit exists under the following conditions:

(a) None of the eigenvalues λ for matrix A exceeds 1 in absolute value.

(b) If $\lambda = 1$, then it is the only eigenvalue, because the stochastic matrix A will have only single blocks in the canonical Jordan form.

Therefore, there are two possible options:

• Beginning with some h, the elements of matrix A^h do not change by more than the given small error and, therefore, the solution is found.

• Matrix A^h will change with the given period. Then it is necessary to find the period and the average between all variants of A^h falling into it.

In this study, the informational structure and software implementation of factor modeling algorithms were developed. The software implementation of the calculation of the influence coefficients in the factor–frame model was developed in Ruby. JSON, the text format of the data exchange based on JavaScript, was used as the input data format [9]. This made it possible to concisely describe the required frames, factors, and relations between them. The software implementation is presented at ws-dss.com.

3. APPLICATION OF THE METHOD TO EVALUATE THE DEVELOPMENT STRATEGIES

According to the theory of M. Porter, successful competitive development is modeled base on the factors of production, investment, innovation, and welfare [10]. At the enterprise level, innovation activity is defined as the process of creating and formalizing innovations (inventions, patents, knowhow, management or production technologies, etc.), as well as introducing or expanding them in order to obtain economic, social, scientific, technological, environmental, or other types of effects [11].

Figure 2 shows the frame-factor model of an innovative enterprise. Frames are marked by rectangles. The name of the frame is given in the upper section of the rectangle. The modeled factors are successfully structured by frame slots. They are shown in the lower section of the frames. The influence of the frames on each other is shown by the black arrows.

The enterprise can operate in several markets and produce several types of products. If they are added to the model as independent frames with the corresponding slots and influence relations, the visual perception of the model will be significantly hampered. All products and all markets have the same features. Therefore, the concept of generalization shown by the white arrow is used to model them. A similar diagram makes it possible to simplify the variations of the model's structure, e.g., upon the occurrence of new

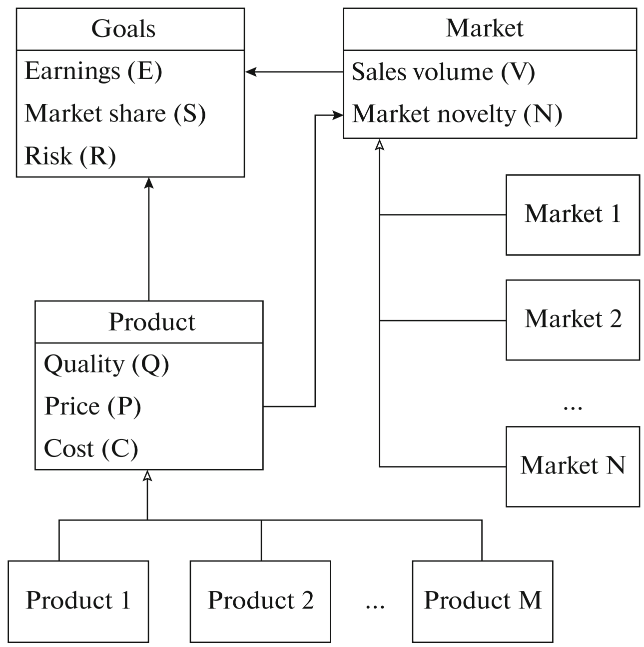

Fig. 2. Frames of an innovative enterprise.

types of products and in the analysis of the expediency of developing new markets. Within the model we can carry out "what if" analysis, e.g., what will happen to earnings if we bring to the market a product with certain characteristics.

Table 1 shows the calculated values of the elements of matrix A for the problem of choosing a strategy for entering one of the alternative markets (Market1 or Market2) provided that one product (Product1) is used. These coefficients were obtained by applying the procedure of expert pairwise comparisons and calculation of eigenvalues according to formulas (3) and (4). The following abbreviations were used to denote the factors: E is the earnings, I is the income, R is the risk, V is the sales volume, N is the market novelty, Q is the quality, P is the price, and C is the cost.

The degree of influence of slots of one frame on one slot of another frame is normalized, e.g., for the influence on E: V + N = 1 and Q + P + C = 1.

As a result of the application of the calculation procedure by formulas (7) and (8), matrix A^2 was obtained, which remains unchanged up to an error of $\varepsilon = 2.886\text{E}{-}03$. It is shown in Table 2. The zero val-

Table 1. Initial values of coefficients of influence of factors

	E	I	R	V	N	Market1	Market2	Q	P	C	Product1
E	0	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	0	0	0
R	0	0	0	0	0	0	0	0	0	0	0
V	0.83	0.5	0.13	0	0	0	0	0	0	0	0
N	0.17	0.5	0.88	0	0	0	0	0	0	0	0
Market1	0	0	0	0.7	0.7	1.0	0	0	0	0	0
Market2	0	0	0	0.3	0	0	1.0	0	0	0	0
Q	0.14	0.29	0.33	0.33	0.6	0	0	0	0	0	0
P	0.5	0.57	0.26	0.33	0.28	0	0	0	0	0	0
C	0.36	0.14	0.41	0.33	0.13	0	0	0	0	0	0
Product1	0	0	0	0	0	0	0	1.0	0.5	0.3	1.0

Table 2. Limit values of coefficients of influence of factors

	E	I	R	V	N	Market1	Market2	Q	P	C	Product1
E	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
I	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
R	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
V	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
N	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Market1	**0.18**	**0.19**	**0.2**	0.35	0.41	1.0	0.0	0.0	0.0	0.0	0.0
Market2	**0.06**	**0.04**	**0.01**	0.15	0.0	0.0	1.0	0.0	0.0	0.0	0.0
Q	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
P	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Product1	0.76	0.77	0.79	0.5	0.59	0.0	0.0	1.0	1.0	1.0	1.0

ues in matrix A^2 suggest that the influence of one factor on the corresponding other factor disappears over time and extends to other factors.

As can be seen from Table 2, it is more reasonable to enter market 1, because this solution is Pareto optimal for all three target factors. If the Pareto optimal solution was not obtained, it would have been necessary to introduce the integral factor of the enterprise's performance or use one of the methods of the multicriteria analysis of the alternatives [12—14].

4. APPLICATION OF THE METHOD TO SELECT AN ELECTRONIC FLIGHT BAG FOR THE CREW

Another example for approbation of the developed mathematical apparatus is the problem of selecting an electronic flight bag (EFB) for flight crews. This problem was formulated in [15]. The following parameters were used as factors to evaluate the usefulness of the selection: Cost, Model Novelty, and Ergonomics. These factors are influenced by the selection of the EFB model: iPad Air, iPad Pro, and iPad Mini. The selection factors are Frame 0. The EFB type is Frame 1. In this problem, the model had two frames with three slots in each.

The degree of influence of the EFB frame on the slot Cost was set by the following pairwise comparisons:

$$U_{11} = \begin{Vmatrix} 1 & 7 & 4 \\ 1/7 & 1 & 1/5 \\ 1/4 & 5 & 1 \end{Vmatrix}.$$

The maximum eigenvalue was $\lambda_{11} = 3.12$. Its corresponding eigenvector was $\mathbf{w}_{11} = (0.94, 0.09, 0.33)$. The other eigenvectors of influence were defined in the same way:

$\mathbf{w}_{12} = (-0.12, -0.91, -0.4)$ is the vector of influence of the EFB on the slot Novelty.

$\mathbf{w}_{13} = (-0.35, -0.2, -0.92)$ is the vector of influence of the EFB on the slot Ergonomics.

$\mathbf{w}_{24} = (0.85, 0.26, 0.47)$ is the vector of influence of the criteria on the slot iPad Air.

$\mathbf{w}_{25} = (-0.85, -0.19, -0.49)$ is the vector of influence of the criteria on the slot iPad Pro.

$\mathbf{w}_{26} = (0.71, 0.24, 0.71)$ is the vector of influence of the criteria on the slot iPad Mini.

After normalization, the matrix of the mutual influence of all the factors was constructed:

$$A^* = \begin{vmatrix} 0.5 & 0.0 & 0.0 & 0.27 & 0.28 & 0.21 \\ 0.0 & 0.5 & 0.0 & 0.08 & 0.06 & 0.07 \\ 0.0 & 0.0 & 0.5 & 0.15 & 0.16 & 0.21 \\ 0.34 & 0.04 & 0.12 & 0.5 & 0.0 & 0.0 \\ 0.03 & 0.32 & 0.07 & 0.0 & 0.5 & 0.0 \\ 0.12 & 0.14 & 0.31 & 0.0 & 0.0 & 0.5 \end{vmatrix}.$$

After six iterations, the matrix elements stopped changing significantly (error e = 1.077E–14). The matrix took the following form:

$$A^6 = \begin{vmatrix} 0.25 & 0.25 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0.07 & 0.07 & 0.07 & 0.07 & 0.07 & 0.07 \\ 0.18 & 0.18 & 0.18 & 0.18 & 0.18 & 0.18 \\ 0.22 & 0.22 & 0.22 & 0.22 & 0.22 & 0.22 \\ 0.09 & 0.09 & 0.09 & 0.09 & 0.09 & 0.09 \\ 0.19 & 0.19 & 0.19 & 0.19 & 0.19 & 0.19 \end{vmatrix}.$$

Therefore, the values of the factors corresponding to the generalized priorities of the EFBs are as follows:

- iPad Air, 0.22;
- iPad Pro, 0.09;
- iPad Mini, 0.19.

The EFB with the maximum factor value should be selected for the flight crews, i.e., iPad Air.

5. CONCLUSIONS

The developed model can be successfully used to solve the following problems:

- Identifying the factors that influence the performance of the analyzed system the most.
- Forecasting the development of complex dynamic systems with feedback.
- Evaluating and taking management decisions based on the analysis of the factors that are the criteria in the considered problems.
- Modeling the innovative activity of enterprises under the conditions of weakly structured, incomplete, and fuzzy information.

This approach can be successfully applied to various high-dimensional factor models in R&D and socioeconomic problems, as well as in modeling the development of the economy of corporations, macroregions, and states.

REFERENCES

1. Jae-On Kim and C. W. Mueller, *Factor Analysis: Statistical Methods and Practical Issues* (SAGE Publications, Thousand Oaks, CA, 1978).
https://doi.org/10.4135/9781412984256
2. B. N. Chetverushkin, V. P. Osipov, and V. I. Baluta, "Approaches to modeling the consequences of decision-making in the context of opposition," Prepr. Inst. Prikl. Mat. im. M. V. Keldysha, No. 43 (2018).
https://doi.org/10.20948/prepr-2018-43
3. M. Minsky, "A framework for representing knowledge," in *The Psychology of Computer Vision,* Ed. by P. H. Winston (McGraw-Hill, New York, 1975), pp. 163–189.
https://doi.org/10.1016/B978-1-4832-1446-7.50018-2
4. R. Sedgewick, *Algorithms in C++ Part 5: Graph Algorithms,* 3rd ed. (Addison-Wesley, Princeton, 2002).
5. T. L. Saaty and L. G. Vargas, *Models, Methods, Concepts and Applications of the Analytic Hierarchy Process* (Kluwer Academic Publishers, Boston, 2000).
6. T. L. Saaty, *Decision Making With Dependence and Feedback: The Analytic Network Process* (RWS Publications, New York, 2001).
7. I. I. Volkov, "Cesàro summation methods," in *Encyclopedia of Mathematics,* Ed. by M. Hazewinkel, (Springer/Kluwer Academic Publishers, 2001).

8. G. H. Golub and C. F. Van Loan, *Matrix Computations,* 3rd ed. (Johns Hopkins University Press, Baltimore, 1996).
9. *Information Technology — The JSON Data Interchange Syntax: ISO/IEC 21778:2017* (Int. Org. Stand., 2017).
10. P. Porter, *The Competitive Advantage of Nations* (Macmillan, London, 1990).
11. R. A. Karaev, I. I. Safarli, M. A. Nagiev, T. F. Abduragimov, and R. G. Gyul'mamedov, "Cognitive analysis and management of innovative projects of enterprises," Inf.-Upr. Sist., No. 4, 63–68 (2010).
12. V. P. Osipov and V. A. Sudakov, "A combined method for multi-criteria decision making support," Prepr. Inst. Prikl. Mat. im. M. V. Keldysha, No. 30 (2015).
13. V. P. Osipov and V. A. Sudakov, "Multicriteria decision analysis with fuzzy preference areas," Prepr. Inst. Prikl. Mat. im. M. V. Keldysha, No. 6 (2017).
https://doi.org/10.20948/prepr-2017-6
14. V. A. Nesterov, B. V. Obnosov, and V. A. Sudakov, "Multi-criteria assessment of military equipment using a hybrid preference function on the example of unmanned aerial vehicles," Vooruzhenie Ekon. **33** (4), 55–66 (2015).
15. A. V. Dutov, V. A. Nesterov, V. A. Sudakov, and K. I. Sypalo, "Fuzzy preference domains and their use for selecting an electronic flight bag for flight crews," J. Comput. Syst. Sci. Int. **57**, 230–238 (2018).

Translated by O. Pismenov