



PROJECT ESTIMATION OF DEVELOPMENT OF COMPLEX TECHNICAL SYSTEMS AT THE DESIGNING STAGE UNDER CONDITIONS OF INTERVAL UNCERTAINTY USING GENETIC ALGORITHM

A. E. Kolodenkova¹, V. V. Korobkin² and E. A. Khalikova¹

¹Ufa State Aviation Technical University (USATU), Ufa Ul K. Marksa, Russia

²Research Institute of Multiprocessor Computation Systems N.A. A.V. Kalyaev, Taganrog, Ul. Chekhova, Russia

E-Mail: akolodenkova@list.ru

ABSTRACT

Existing R & D advancements are accompanied by increase in complexity of developed engineering systems and facilities, involved in major critical industries, such as nuclear plant industry, oil and gas production, aircraft and defense industry, transport, and so on. At current stage the development of complex engineering systems (CES) is a labor consuming process, since it is related with significant expenditures of material and financial resources under risks stipulated by uncertainty of internal and external factors of project environment, it involves scientific researches, designing and experimental activities, field tests, and development of documentation. Increase in CES complexity results in increased designing times, decreased quality of CES, violation of CES designing terms, or just in project failure. Thus, high quality designing of CES under conditions of uncertainty is an urgent and significant problem for numerous modern enterprises. This work, aiming at elimination of risks, decrease in managerial errors adopted by project directors under conditions of interval uncertainty of initial data, proposes to apply estimation of time and cost of CES designing on the basis of genetic algorithm comprised of analysis of possible designing alternatives.

Keywords: genetic algorithm, designing of complex engineering systems, interval uncertainty.

INTRODUCTION

Designing of CES is one of the most important and initial stages of product life cycle. It is comprised of activities aimed at development of feasibility study, researching activities; studies of sales markets and consumer requirements, operation conditions, material resources; formation of requirements to level and quality of new (retrofitted, modified, improved) product; tender for development and/or production; implementation of engineering solutions into specifications aimed at the most cost efficient operations, testing of specimens [3, 5, 9].

Designing stage is a complicated, multicriterion, inertial process characterized by risks, uncertainty of initial data and performed by large crews of executors [1-2].

Despite numerous works in Russia and abroad, the estimation problem of CES development project at designing stage is still unsolved [4, 6-8, 13-19].

This work proposes to apply estimation of time and cost of CES designing on the basis of genetic algorithm under conditions of interval uncertainty of initial data.

ESTIMATION PROCEDURE OF COST AND TIME OF DESIGNING OF SYSTEMS USING GENETIC ALGORITHM

The concept of estimation of CES projects at designing stage using our modified genetic algorithm is comprised of analysis of possible alternatives of project development. Estimation of each k -th alternative of project development is based on expected cost of alternative $s(x^{(k)})$, $k = 1, \dots, m$, expected development time of alternative $t(x^{(k)})$, $k = 1, \dots, m$ and generation of multicriterion estimation $P(x^{(k)})$.

Now let us describe the adopted assumptions:

a) four developed alternatives of project development are analyzed ($x^{(1)}, x^{(2)}, x^{(3)}, x^{(4)}$). Numerous alternatives lead to high validity of adopted managerial decisions at designing stage. However, consideration of excessive number of alternatives increases efforts, time consumptions and confusions. Thus, project managers usually consider only few alternatives.

b) four partial criteria are considered, which characterize alternatives and can be presented as interval and fuzzy triangle and trapezoidal numbers, their boundaries can intersect [12, 20];

c) partial criterion of alternatives development cost is denoted as $s(x^{(k)})$ and can be presented as follows:

$s(x^{(k)}) = [s_1^{(k)}, s_2^{(k)}]$, where $s_1^{(k)}$, $s_2^{(k)}$ is the minimum and maximum development cost of alternatives;

d) partial criterion of alternatives development time is denoted as $t(x^{(k)})$ and can be presented as follows:

$t(x^{(k)}) = [t_1^{(k)}, t_2^{(k)}]$, where $t_1^{(k)}$, $t_2^{(k)}$ is the minimum and maximum designing time of systems.

Formulation of the problem. There is a limited set of alternatives of project development $X = \{x^{(1)}, x^{(2)}, \dots, x^{(m)}\}$, where each alternative $x^{(k)} \in X$, $k = \overline{1, m}$ is estimated by a sequence of n non-normalized interval partial criteria $C = \langle c_j(x^{(k)}) \rangle$,

$j = \overline{1, n}$. It is required to determine the vector of expected cost and time of development for each k -th alternative of project implementation: $s(x^{(k)})$,



$t(x^{(k)}), k = \overline{1, m}$, as applied to estimation of generalized usefulness characterizing feasibility of each k -th alternative of project development at CES designing stage:

$$P(x^{(k)}) = \sum_{j=1}^n w_j^{\text{nor}} p_j^{\text{nor}}(x^{(k)}), k = \overline{1, m}, j = \overline{1, n}, (1)$$

with possibility of subsequent selection of acceptable alternative.

Here w_j^{nor} is the normalized interval coefficient of relative importance of j -th partial criterion of alternatives $x^{(k)} \in X$; $p_j^{\text{nor}}(x^{(k)})$ are the normalized interval partial criteria of alternatives $x^{(k)} \in X$ ($0 < p_j^{\text{HOP}}(x^{(k)}) < 1$).

Therefore, the problem of feasibility estimation of alternatives of project development of CES at designing stage is described by the following target function:

$$P(x^{(k)}) = \sum_{j=1}^n w_j^{\text{nor}} p_j^{\text{nor}}(x^{(k)}) \rightarrow \max, x^{(k)} \in X, (2)$$

and limitations:

$$s_{\min} \leq s(x^{(k)}) \leq s_{\max}, (3)$$

$$t_{\min} \leq t(x^{(k)}) \leq t_{\max}. (4)$$

The considered problem is of combinatory type and can be formulated as searching for the best solution. In order to determine the expected cost $s(x^{(k)})$ and designing time of alternative $t(x^{(k)})$ it is proposed to apply genetic algorithm comprised of sixteen steps (Figure-1).

- Step 1.** Input of initial data: $c_j(x^{(k)})$ - non-normalized interval partial criteria of alternatives; s_{\min} and s_{\max} - minimum and maximum costs of alternatives; t_{\min} and t_{\max} - minimum and maximum alternatives development time; kol - number of iterations; d - crossover position; q - mutation position; w_j - interval coefficient of relative importance of the j -th partial criterion of alternatives.
- Step 2.** Determination of required optimization conditions: target function (2) and limitations (3)-(4).
- Step 3.** Generation of initial population – parent objects (four variants of cost vectors and development time for each k -th alternative of project implementation) using random number generator:

$$s_1^P(x^{(1)}) = (s_1^{P(1)}, s_2^{P(1)}, s_3^{P(1)}, s_4^{P(1)}),$$

$$s_2^P(x^{(2)}) = (s_1^{P(2)}, s_2^{P(2)}, s_3^{P(2)}, s_4^{P(2)}),$$

$$s_3^P(x^{(3)}) = (s_1^{P(3)}, s_2^{P(3)}, s_3^{P(3)}, s_4^{P(3)}),$$

$$s_4^P(x^{(4)}) = (s_1^{P(4)}, s_2^{P(4)}, s_3^{P(4)}, s_4^{P(4)}).$$

$$t_1^P(x^{(1)}) = (t_1^{P(1)}, t_2^{P(1)}, t_3^{P(1)}, t_4^{P(1)}),$$

$$t_2^P(x^{(2)}) = (t_1^{P(2)}, t_2^{P(2)}, t_3^{P(2)}, t_4^{P(2)}),$$

$$t_3^P(x^{(3)}) = (t_1^{P(3)}, t_2^{P(3)}, t_3^{P(3)}, t_4^{P(3)}),$$

$$t_4^P(x^{(4)}) = (t_1^{P(4)}, t_2^{P(4)}, t_3^{P(4)}, t_4^{P(4)}),$$

where P is the parent object.

Step 4. If the terms (3) and (4) are valid for parent objects, then go to Step 5; if the terms (3) and (4) are not valid, then go to Step 3.

Step 5. Addition of vectors $s_k^P(x^{(k)})$ and $t_k^P(x^{(k)})$ into the population.

Step 6. One-point crossover with regard to the obtained variants of cost vectors of the k -th alternatives $s^P(x^{(k)})$ (crossover point $d < 4$). As a result of the crossover 12 child objects are obtained:

$$s_1^\Pi(x^{(1)}) = (s_1^{P(1)}, s_d^{P(1)}, s_{d+1}^{P(2)}, s_4^{P(2)}),$$

$$s_2^\Pi(x^{(2)}) = (s_1^{P(2)}, s_d^{P(2)}, s_{d+1}^{P(1)}, s_4^{P(1)}),$$

$$s_3^\Pi(x^{(1)}) = (s_1^{P(1)}, s_d^{P(1)}, s_{d+1}^{P(3)}, s_4^{P(3)}),$$

$$s_4^\Pi(x^{(3)}) = (s_1^{P(3)}, s_d^{P(3)}, s_{d+1}^{P(1)}, s_4^{P(1)}),$$

$$s_5^\Pi(x^{(1)}) = (s_1^{P(1)}, s_d^{P(1)}, s_{d+1}^{P(4)}, s_4^{P(4)}),$$

$$s_6^\Pi(x^{(4)}) = (s_1^{P(4)}, s_d^{P(4)}, s_{d+1}^{P(1)}, s_4^{P(1)}),$$

$$s_7^\Pi(x^{(2)}) = (s_1^{P(2)}, s_d^{P(2)}, s_{d+1}^{P(3)}, s_4^{P(3)}),$$

$$s_8^\Pi(x^{(3)}) = (s_1^{P(3)}, s_d^{P(3)}, s_{d+1}^{P(2)}, s_4^{P(2)}),$$

$$s_9^\Pi(x^{(2)}) = (s_1^{P(2)}, s_d^{P(2)}, s_{d+1}^{P(4)}, s_4^{P(4)}),$$

$$s_{10}^\Pi(x^{(4)}) = (s_1^{P(4)}, s_d^{P(4)}, s_{d+1}^{P(2)}, s_4^{P(2)}),$$

$$s_{11}^\Pi(x^{(3)}) = (s_1^{P(3)}, s_d^{P(3)}, s_{d+1}^{P(4)}, s_4^{P(4)}),$$

$$s_{12}^\Pi(x^{(4)}) = (s_1^{P(4)}, s_d^{P(4)}, s_{d+1}^{P(3)}, s_4^{P(3)}),$$

where Π is the child object.

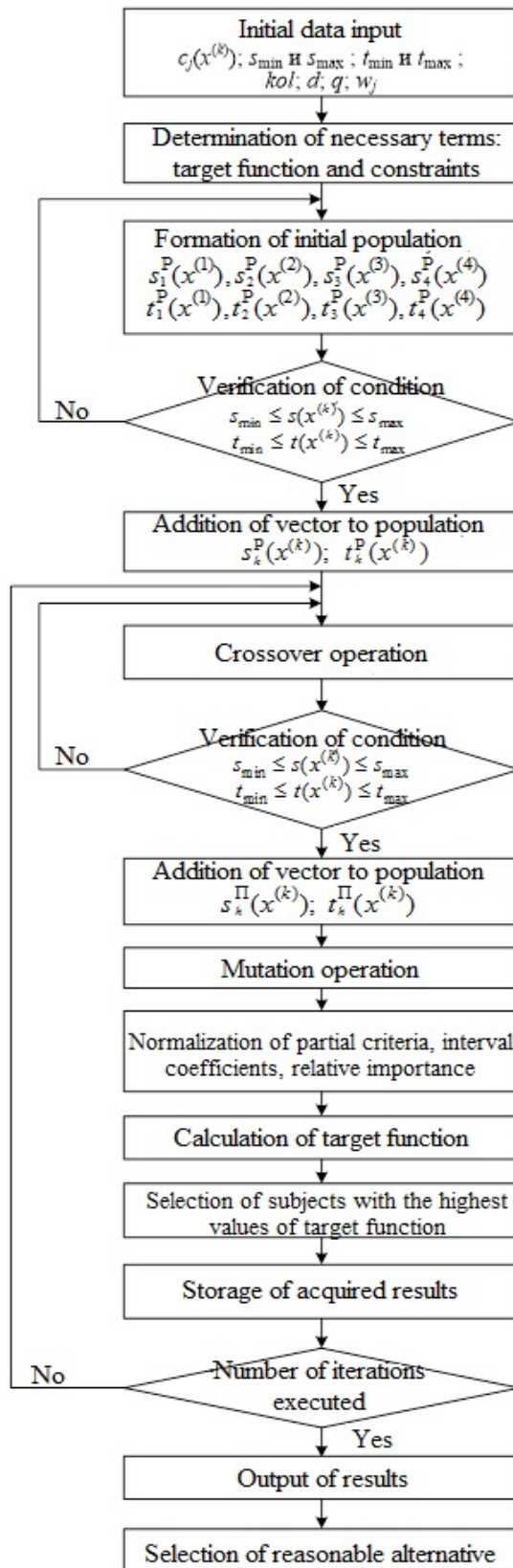


Figure-1. Genetic algorithm of determination of expected time and cost of alternative.

One-point crossover with regard to the obtained variants of time vectors of development of the k -th alternative $t^P(x^{(k)})$ (crossover point $d < 4$) is performed similarly. After crossover operation 12 child objects are obtained.

The crossover operator makes it possible on the basis of crossing of parent chromosomes to create child chromosomes. One-point crossing is comprised of cutting of parent chromosomes in randomly selected common cutting point and exchanging with right-hand parts (chromosome tailings).

Step 7. If the terms (3) and (4) are valid for the obtained child objects, then go to Step 8; if the terms (3) and (4) are not valid, then go to Step 6.

Step 8. Addition of vectors $s_h^{\Pi}(x^{(k)})$ and $t_h^{\Pi}(x^{(k)})$, $h = 1, \dots, 12$ into the population.

Step 9. One-point mutation of the obtained 12 child objects (mutation point $q < 4$). The child object obtained by mutation provides new child object with mutated genes $\bar{s}_1^M(x^{(1)}) = (s_1^M, s_q^M, \bar{s}_{q+1}^M, s_4^M)$. This operation is applied similarly to the other child objects.

One-point mutation of the obtained 12 child objects (mutation point $q < 4$) for alternative development time is performed in a similar way.

Mutation operator makes it possible to add variations into objects which subsequently obtain new properties. One-point mutation is comprised of random selection of gene, which exchanges its value with neighbor gene.

Step 10. Normalizing partial criteria $p_j^{\text{nor}}(x^{(k)})$, including expected cost and expected time of development for each k -th alternative of project development [10, 11] is performed as follows:

a) for intervals

$$c_j(x^{(k)}) = [c_{j1}(x^{(k)}), c_{j2}(x^{(k)})]$$

$$p_j^{\text{nor}}(x^{(k)}) = \left[\frac{c_{j1}(x^{(k)})}{c_j^{\max}}, \frac{c_{j2}(x^{(k)})}{c_j^{\max}} \right],$$

$$c_j^{\max} = \max_{1 \leq k \leq m} \{c_{j2}^{(k)}\},$$

where $c_{j1}(x^{(k)})$, $c_{j2}(x^{(k)})$ are the minimum and maximum values of the interval;

b) for fuzzy triangle number

$$c_j(x^{(k)}) = [c_{j1}(x^{(k)}), c_{j2}(x^{(k)}), c_{j3}(x^{(k)})]$$



$$p_j^{\text{nor}}(x^{(k)}) = \left[\frac{c_{j1}(x^{(k)})}{c_j^{\text{max}}}, \frac{c_{j2}(x^{(k)})}{c_j^{\text{max}}}, \frac{c_{j3}(x^{(k)})}{c_j^{\text{max}}} \right],$$

$$c_j^{\text{max}} = \max_{1 \leq k \leq m} \{c_{j3}^{(k)}\},$$

where $c_{j1}(x^{(k)})$, $c_{j2}(x^{(k)})$, $c_{j3}(x^{(k)})$ are the minimum, the most expected, and the maximum values of the interval, respectively;

c) for fuzzy trapezoid numbers

$$c_j(x^{(k)}) = [c_{j1}(x^{(k)}), c_{j2}(x^{(k)}), c_{j3}(x^{(k)}), c_{j4}(x^{(k)})]$$

$$p_j^{\text{nor}}(x^{(k)}) = \left[\frac{c_{j1}(x^{(k)})}{c_j^{\text{max}}}, \frac{c_{j2}(x^{(k)})}{c_j^{\text{max}}}, \frac{c_{j3}(x^{(k)})}{c_j^{\text{max}}}, \frac{c_{j4}(x^{(k)})}{c_j^{\text{max}}} \right]$$

$$, c_j^{\text{max}} = \max_{1 \leq k \leq m} \{c_{j4}^{(k)}\},$$

where $c_{j1}(x^{(k)})$, $c_{j4}(x^{(k)})$ are the pessimistic and the optimistic estimations of interval boundaries, $[c_{j2}(x^{(k)}), c_{j3}(x^{(k)})]$ is the interval of the most expected values.

It should be mentioned that normalization of partial criteria is necessary, since their numerical values are described by different units and orders, which complicates further operations with these criteria.

Normalization of interval coefficient of relative importance of the j -th partial criteria of alternatives is performed as follows:

$$w_j^{\text{nor}} = \frac{w_j}{\sum_{j=1}^n w_j}, \quad j = \overline{1, n},$$

where w_j is the interval coefficient of relative importance of the j -th partial criterion, which can be generated on the basis of opinion survey of project executors and presented in the form of interval, fuzzy triangle and trapezoid numbers.

If the normalized interval coefficient of relative importance of the j -th partial criterion is in the form of interval $w_j^{\text{nor}} = [\alpha_{j1}, \alpha_{j2}]$, where α_{j1} , α_{j2} are the minimum and the maximum values of interval, then it is

assumed that $\sum_{j=1}^n \alpha_{j1} < 1$, $\sum_{j=1}^n \alpha_{j2} > 1$, since

otherwise the Eq. (2) has no solution due to impossibility

of limitation $\sum_{j=1}^n \alpha_j = 1$.

If the normalized interval coefficient of relative importance of the j -th partial criterion is in the form of fuzzy triangle number $w_j^{\text{nor}} = [\alpha_{j1}, \alpha_{j2}, \alpha_{j3}]$, where α_{j1} , α_{j2} , α_{j3} are the minimum, the most expected,

and the maximum values of interval, then it is assumed

that $\sum_{j=1}^n \alpha_{j1} < 1$, $\sum_{j=1}^n \alpha_{j3} > 1$. Otherwise the Eq. (2)

has no solution.

If the normalized interval coefficient of relative importance of the j -th partial criterion is in the form of fuzzy triangle number $w_j^{\text{nor}} = [\alpha_{j1}, \alpha_{j2}, \alpha_{j3}, \alpha_{j4}]$,

where α_{j1} , α_{j4} are the pessimistic and optimistic

estimations of interval boundaries, $[\alpha_{j2}, \alpha_{j3}]$ is the

interval of the most expected values, then it is assumed

that $\sum_{j=1}^n \alpha_{j1} < 1$, $\sum_{j=1}^n \alpha_{j4} > 1$. Otherwise the Eq. (2)

has no solution.

Step 11. Calculation of values of target function (1) for all objects of the population, that is, for all variants of vectors of expected costs and development time of project alternatives.

Step 12. Selection of four objects (four vectors) of 12 available with the highest values of target function, which will be the parents for the next iteration (generation), or, after all iterations, the final results of calculations.

Since four alternatives of project development are considered in the proposed problem, then only four vectors of expected cost and development time are selected. The objects are selected on the basis of ranking, that is, the population objects are ranked according to the values of their fitness function (ascending ranking). Since the estimations of generalized usefulness $P(x^{(k)})$ are presented in the form of fuzzy trapezoid numbers, then acceptable project alternative is selected using the Chew–Park method.

The Chew–Park method is as follows: at first the parameter w is fixed, then each trapezoid number $X = [x_1, x_2, x_3, x_4]$ is associated with (real) number

$$cp(X) = \frac{x_1 + x_2 + x_3 + x_4}{4} + w \frac{x_2 + x_3}{2}.$$

It should be mentioned that there is no unified approach to select the parameter w , thus, by default $w = 1$. Sorting is performed in ascending order of $cp(X)$.

Step 13. Storage of the obtained results.

Step 14. If the number of iterations is met, then go to Step 14; if the condition is not met, then go to Step 6.

Step 15. Output of results (vectors of expected cost $s(x^{(k)})$, expected development time of alternatives $t(x^{(k)})$ and multi-criterion estimation $Cp(P(x^{(k)}))$ for each k -th alternative of project development).

Step 16. Selection of acceptable alternative. Acceptable alternative is that with the highest generalized usefulness.



End of algorithm.

Now let us consider an example. The calculations are aimed at searching for optimum solution of feasibility estimation of alternative at designing stage in order to perform its subsequent implementation. Let us consider operation of the above described genetic algorithms using the following example.

EXAMPLE

There are four alternatives ($x^{(1)}, x^{(2)}, x^{(3)}, x^{(4)}$) of project development, which are described by four partial

criteria: c_1 - the time of project implementation, months; c_2 - the number of executors, persons.; c_3 - the probability of project success with occurring situations preventing partially or completely achievement of project targets, as well as approximate cost of the project $s(x_i) = [600, 760]$ thousand rubles, $t(x_i) = [6, 9]$ months ($i = 1, \dots, 4$). It is required to estimate the alternatives with subsequent selection of alternative of project development.

The calculated values of partial criteria for each alternative are summarized in Table-1.

Table-1. Calculated values of partial criteria of implemented alternatives.

Projects alternatives	Partial criteria of feasibility of alternatives		
	c_1	c_2	c_3
$x^{(1)}$	[9.5, 11]	[10, 15]	[0.8, 0.84]
$x^{(2)}$	[8.7, 10]	[8, 12]	[0.83, 0.87]
$x^{(3)}$	[9.2, 10.3]	[9, 12]	[0.8, 0.85]
$x^{(4)}$	[9.5, 10.7]	[10, 13]	[0.84, 0.9]

On the basis of survey of executor opinions, the coefficients of relative importance of partial criteria were obtained in the form of trapezoid numbers:

$$w_1 = [3, 4], w_2 = [3, 4], \\ w_3 = [2, 3], w_4 = [2, 5, 3, 2],$$

Initial values of genetic algorithm are as follows: crossover point - 1; mutation point - 1; number of iterations - 10.

Table-2 summarizes the test results of the developed software characterized by capability to operate with interval values.

Table-2. Results of optimization.

Project alternatives	Project cost vector	Designing time vector	Estimation of generalized usefulness
$x^{(1)}$	[600, 721]	[6, 7]	0.94
$x^{(2)}$	[600, 736]	[6, 7.5]	0.72
$x^{(3)}$	[600, 709]	[6, 8]	0.55
$x^{(4)}$	[600, 750]	[6, 8.5]	0.69

It follows from Table 2 that the alternative x_1 with costs of [600, 721] thousand rubles, development time of [6, 7] months and respective generalized usefulness $F(x^{(1)}) = 0.94$ is the best alternative of CES development at designing stage.

CONCLUSIONS

A method of CES project estimation at designing stage has been developed on the basis of modified genetic algorithm, comprised of determination of vector of expected cost and time for each alternative of project development with regard to estimation of generalized usefulness characterizing feasibility of each alternative. The proposed method makes it possible to save financial expenses, to eliminate risks, to reduce managerial errors of

project managers under conditions of interval initial data, to select the most acceptable alternative of CES development at designing stage when the values of partial criteria are intersected, to adopt decision either to continue the project or to generate new alternatives, that is, to adopt weighted decision about its continuation.

ACKNOWLEDGEMENTS

The work was supported by Russian Foundation for Basic Research, project No. 17-08-00402, System of methods and models of intellectual in-line control and prevention of hazardous situations at the stage of design of complex engineering systems for facilities of critical infrastructure.



REFERENCES

- [1] Arkhipenkov S. Ya. 2009. Lectures on Management of Software Projects. Moscow, Retrieved 14.01.12017 from http://www.arkhipenkov.ru/resources/sw_project_management.pdf.
- [2] Boehm B. (n. d.) Software development cost estimation approaches - a survey. Retrieved 14.01.12017 from <http://csse.usc.edu/TECHRPTS/2000/usccse2000-505/usccse2000-505.pdf>.
- [3] Chuikin A. M. 2000. Development of Managerial Decisions; Guidebook. Kaliningrad, Kaliningrad University.
- [4] DeMarco T. and Lister T. 2003. Waltzing with Bears. Managing Risks on Software Projects. Published by Dorset House.
- [5] Dodonov V. A. (n. d.). Methodology of System Designing and development of Automated Systems of Data Processing and Management; <https://refdb.ru/look/2835012.html>.
- [6] Faheem A., Salah B., Adel S. and Issa K. (n. d.). Integrating function point project information for improving the accuracy of effort estimation. Retrieved 14.01.12017 from http://faculty.uaeu.ac.ae/ikhalil/docs/SWE_FP.pdf.
- [7] Gharehchopogh F. S. 2015. A new approach based on continuous genetic algorithm in software cost estimation. Journal of Scientific Research and Development. 2(4): 87-94.
- [8] Gladkov L. A. (n. d.). Solutions of Managerial and Optimization Tasks on the Basis of Hybrid Intelligent Methods. Retrieved 14.01.12017 from <http://refdb.ru/look/2068541.html>.
- [9] ISO/IEC/IEEE 24765-2010 Systems and software engineering - Vocabulary. JTC 1/SC 7, (2010), p. 418.
- [10] Kolodenkova A. E. 2015. Assessment of Feasibility of Data Management Systems of Nuclear Plants upon Interval Pattern of Initial Data. Izv. VUZov. North Caucasian Region. Technical Sciences. 4(185): 28-34.
- [11] Krioni N.K., Kolodenkova A.E., Korobkin V.V. and Khalikova E.A. 2016. Estimation of time expenditures for software development of data management systems with fuzzy initial data. Vestn. komput. inform. Tekhnolog. 1(139): 10-16.
- [12] Levin V. I. 1998. Comparison of interval values and optimization of indefinite systems. Information Technologies. 7, 22-32.
- [13] Levinson M. 2009. Common errors of project management. Retrieved 14.01.12017 from <https://www.osp.ru/cio/2009/02/5337322/>.
- [14] Lipaev V. V. 2006. Software Engineering. Foundations, Guidebook. Moscow: TEIS.
- [15] McConnel S. 2006. Software Estimation: Demystifying the Black Art (Developer Best Practices). Microsoft Press.
- [16] Orlov A. I. 2003. High statistic technologies. Zavod. Lab. 69(11): 55-60.
- [17] Orlov. S. A. 2002. Technologies of Software development. Saint Petersburg: Piter.
- [18] Standish Group 2015 Chaos Report - Q & A with Jennifer Lynch; Retrieved 14.01.12017 from <http://www.infoq.com/articles/standish-chaos-2015>.
- [19] Vasilev A. S. and Matveikin V. G. 2008. Modification of genetic algorithm for solution of scheduling with constrained resources. Transactions TSTU. 14(2): 310-314.
- [20] Voshchinin A. P. 2002. Interval data analysis: future and challenges. Zavod. Lab. 68(1): 118-126.