

The Optimization Methodology of the Digital Model of the Educational Program Based on Learning-Forgetting Functions

Victor Yu. Stroganov¹ and Yurii B. Tsvetkov^{1*}

¹Bauman Moscow State Technical University, 2nd Baumanskaya str., 5/1, 105005, Moscow, Russia

Abstract. A methodology for curriculum optimizing based on the construction of learning-forgetting terms of educational content is proposed. In the formation of the curriculum, it is important not only to minimize the intervals between educational modules that use the same terms, but also to take into account the forgetting of the studied educational material. The article offers a generalized representation of the optimization methodology of the digital model of the educational program based on modeling the learning-forgetting function of the terms throughout the educational program, which poses the task of optimizing the curriculum in a multi-criteria setting.

1 Introduction

A continuous increase in the requirements for engineering education determines the intensification of the learning process, increasing its effectiveness.

The need to expand the studied subject areas and increase the level of development of educational information imposes increased requirements for the design of appropriate educational programs, and, especially, their key components - curricula [1-8].

Curricula should provide:

- compliance of the set of studied disciplines with the lists of competencies presented in the relevant educational standards;
- correspondence of the time allocated for the study of individual information units (modules, topics), the degree of their use in other disciplines and in the professional activities of the graduate;
- logical harmony in the presentation of educational material, interdisciplinary communications;
- the optimality of the ratio of various types of classroom and independent work of students;
- taking into account the psychophysiological features of the perception of information.

The basis for the formation of curricula in modern conditions is a competency-based approach that involves designing goals, selecting educational content and organizing the educational process based on an assessment of planned learning outcomes. At the same time,

* Corresponding author: tsvetkov@bmstu.ru

attention is focused not on the assimilation of information, but on the student's ability to use it to solve problems, both in the professional field and beyond.

In modern conditions of digital transformation of engineering education, the educational process is becoming the object of mathematical modelling. Models provide a visual representation of the connections between the components of the educational program, identify the structure of its logical construction, establish quantitative and qualitative relationships between elements of the curriculum.

On the basis of models, it is also possible to optimize the curriculum by selecting the content of training, establishing the sequence of study of academic disciplines, designating the relationships and relationships between them.

The analysis of mathematical models of competency-based curricula, developed, inter alia, on the conditions of previous state educational standards, showed that the existing developments are advisory in nature and represent a verbal description of the solutions [1].

As a criterion for the optimality of the curriculum, well-known works often consider the degree of conformity of a plan option with the formulated requirements (limitations), first of all, the dynamics of students' workload.

Building the structure of the graph of the relationship between the sections of the course, identifying the density of information links and dependencies between them - all this is the basis for the analysis of the curriculum in both domestic and foreign practice [1, 4-8], forming the basis for increasing the effectiveness of the educational process [9-18].

It seems important not only to minimize the interval between the periods of studying interrelated sections, but also to take into account the function of forgetting the studied educational material [2, 3].

The aim of the work is a generalized presentation of the optimization methodology of the digital model of the educational program based on the learning – forgetting function.

2 Network model of the curriculum

In the general case, the curriculum is presented in the form of a digraph, where the vertices represent the modules $\mathbf{Mod}=\{Mod_m\}$, $m=1\dots M$ (M is the number of modules), and the arcs $\mathbf{G}=\{g_{ij}\}$, $i,j=1..M$ define their semantic connectivity through terms used in modules. Those, if the term $Term_{m,i}$ is introduced in some Mod_m module, and it is used in another Mod_n module, then it should be in the curriculum of the Mod_m module, which determines the presence of the preceding order of the modules or the corresponding arc. This is fully consistent with the formal network plan model. In the classical layout of modules in the discipline, the graph of the curriculum represents a fairly regular structure (see. Fig. 1).

It should be noted that such a curriculum formation is mainly associated with organizational problems of conducting classes at a university, where there is a certain load of departments, a schedule of teachers with certain working hours, etc.

In this case, the following parameters take place that determine the resource plan of the network plan:

T_{dis} – the duration of reading the discipline (mainly a semester);

V_{dis} – is the total hours allocated to the discipline.

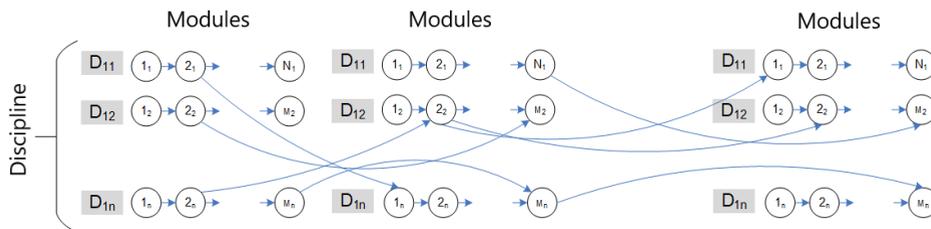


Fig. 1. Graph of curriculum modules grouped into disciplines.

These parameters determine the indicator of the intensity of studying the discipline in the semester, which is equal to $L_{dis} = V_{dis} / T_{dis}$. (e.g. hours per day for a given discipline or hours per week). The student’s total resource is limited by the volume W , equal to 54 hours per week (see. Fig. 2), which sets the task of rational distribution of disciplines by semesters.

If we consider individual plans, it is more likely that a slightly different structure will be rational, which consists of separate modules not integrated into disciplines. Next, we will consider the model of the curriculum only in a modular representation without integration into disciplines.

In this setting, we have a graph in the language of works in which the presence of an arc determines the immediately preceding modules. The module graph in the appendix is weighted by many terms. Those, in the language of network planning, during the implementation of each work (reading the training module), all its terms become active and a program for constructing the functions of learning terms is implemented. If a specific resource is allocated for the work (the student’s time to study this module from the calculation of no more than 54 hours per week), then it does not change until its completion.

Thus, each vertex of the network plan (see Fig. 3) is weighted:

- Implementation time T_{mod} ;
- the amount of V_{mod} resources;
- he set of terms that are used in the module $\{Term_1, Term_2, \dots, Term_n\}$.

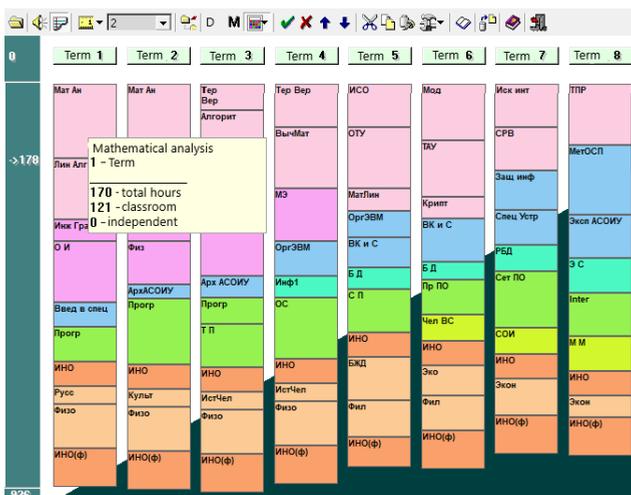


Fig. 2. Resource scheme for the implementation of the network plan of the educational process by semester.

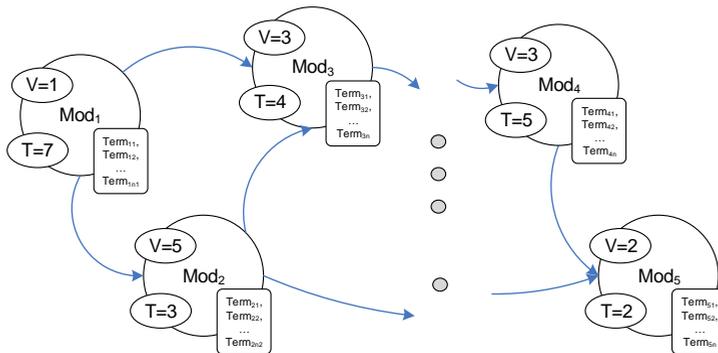


Fig. 3. A digraph of a network plan for studying modules in a language of work.

In the classical model of network planning, there are two types of graph representation. A graph in the language of works and a graph in the language of events.

The network diagram in the language of work is a directed graph whose vertices are the work, and the arcs represent the precedence relation: the arrow is directed from the previous work to the next. Subsequent work can only be started after all previous work has been completed.

The way of presenting the project in the language of work is the most natural, however, the calculation algorithms in this case are more complex than in the language of events. For this reason, historically, the first form of network diagram in the language of events was used. The network graph in the language of events is a directed graph whose vertices are events corresponding to the moments of the beginning and end of work, and arcs indicate work.

In this case, there are recurrent schemes for calculating the values of the early and late terms of reading a module without disrupting the general terms for completing the training program. It also calculates all time reserves for each job, etc. However, when scheduling work in the language of events, there are a number of restrictions [11, 12] which makes it somewhat more difficult to draw up a network plan.

3 Analysis of learning-forgetting patterns

In terms of the formal model of the educational process, in addition to the network planning model, the article solves the problem of modeling the functions of learning to forget terms. They are established as one of the criteria for evaluating the effectiveness of the curriculum. It is assumed that the calculation of learning-forgetting functions of all terms at the end of training is carried out according to the formed curriculum in accordance with the competencies of the specialty in which the curriculum was formed.

There are many analogies to the construction of learning-forgetting functions, namely: models - analogies of physical phenomena and technical systems; theoretical information models; models - analogies of cybernetic systems and others.

The approximation of the “forgetting curves” proposed by G. Ebbinghaus (1885) was based on an exponential function. The explanation of this functional dependence is quite simple - a person has "short-term" and "long-term" memory, characterized by different times of memorization and storage of information (see. Fig. 4).

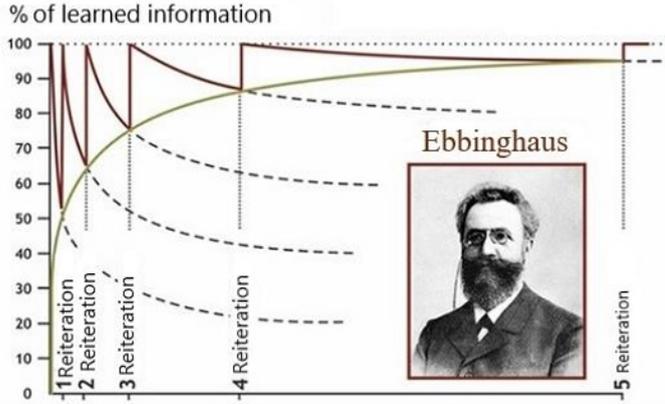


Fig. 4. Ebbinghaus Forgetting Curve.

The processes of iterative learning, as a rule, are characterized by slow-asymptotic learning curves approximated by exponential curves. In general, the exponential curve is described by the dependence.

$$x(t) = x_{\infty} + (x_0 - x_{\infty})e^{-\gamma t}, \quad t > 0, \tag{1}$$

or sequence :

$$x_n = x_{\infty} + (x_0 - x_{\infty})e^{-\gamma n}, \quad n = 0, 1, 2, \dots, m, \tag{2}$$

where t – is the time of learning, n – is the number of iterations (tests, attempts) from the moment of learning (it is assumed that learning begins at time zero), $x(t)$ (x_n) – is the value of the mismatch at time t (at the n -th iterations), x_0 – is the initial value of the mismatch (corresponding to the moment of the beginning of learning), x_{∞} – is the "final" value of the mismatch.

According to the presentation of Shlensky O.F. and Bode B.V. [9] with a constant amount of information arriving per unit of time, an "ideal memory" remembers all the information. In real memory, the amount of information stored per unit time decreases with the growth of already stored information, which is manifested in slow asymptotic behavior. After the end of the learning process, the ideal memory stores information indefinitely, and in real memory the amount of information decreases monotonously (forgetting) after the end of the learning process, and the current forgetting speed is proportional to the amount of information $I(t)$ currently available (see Fig. 5).

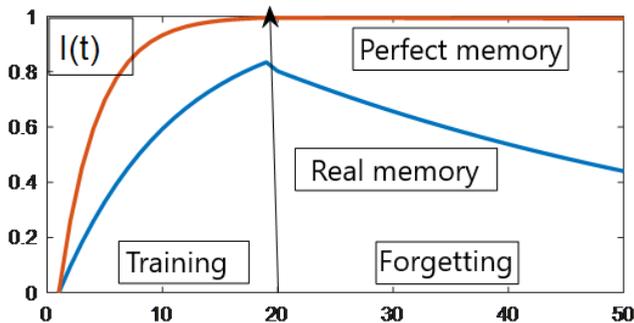


Fig. 5. The construction of learning-forgetting functions according to Shlensky O.F. and Bode B.V.

After the early work of Ebbinghaus and Thorndike, the empirical data obtained during the training began to be expressed in the form of equations that were not based on any theory, but only interpreted the available data. So, A. Shukarev derived the following equation:

$$y = a - b \cdot e^{-c n}, \quad (3)$$

where y – is the assimilation, defined as the number of correct reproductions (successes) per unit of time; n – is the number of tests; a – is the assimilation limit for $n \rightarrow \infty$; b и c – are constants.

T. Robertson proposed an equation obtained by analogy with a monomolecular autocatalytic reaction:

$$y = \frac{b \cdot e^{A n}}{c + e^{A n}}, \quad (4)$$

where y – is assimilation; n – is the number of tests; $A=ab$; a и c – are constants (parameters of the learner); b – is the assimilation limit for $n \rightarrow \infty$.

L. Thurstone [10] proposed the following formula (the so-called hyperbolic law of instruction):

$$y = \frac{a \cdot (n + c)}{(n + c) + b}, \quad (5)$$

where y – is assimilation; n – is the number of tests; a и c – are constants; b – is the speed of learning.

Further development of the application of mathematical methods to the study of problems of learning is associated with the work of K. Hall. He introduced a variable, which he called "the power of skill." It is expressed by the formula:

$$H_R^S = M(1 - e^{-bn}), \quad (6)$$

where H_R^S – is the "strength of skill", or an associative variable linking the stimulus and the response; M – is the asymptotic value of "skill strength»; b – is a parameter expressing the speed of learning; n – is the number of training experiments (or tests with reinforcement). The Hull model allows you to predict the results obtained by training paired associations.

4 Approximation of learning-forgetting functions based on the Laguerre function system

In order to approximate the forgetting function, a certain extension of the exponential dependence is proposed. This dependence is close to the system of Laguerre functions, which represent an orthonormal basis in the space of functions with an integrable square.

$$L_n(x) = L_n^*(x) \cdot e^{-x}, \quad (7)$$

where $L_n^*(x)$ – is an n -th polynomial. This system of functions is convenient for representing exponential functions, Erlang functions, and others.

To fulfill the property of monotonicity of the functions of learning and forgetting, it is necessary that the derivative of the function over the entire admissible region be negative. Moreover, the operations of integration and differentiation do not derive elements from this system of functions. Therefore, a monotonically decreasing function can be obtained by

integrating a function that is negative over the entire domain. This condition is satisfied by the functions inverted to Erlang distributions:

$$E_r(x) = \frac{r\mu(r\mu x)^{r-1} e^{-r\mu x}}{(r-1)!} \tag{8}$$

Since this function represents the distribution density, the integral of it over the entire domain is 1. Therefore, for the forgetting function, the representation:

$$Z(t) = 1 - \int_0^t E_r(x) dx \tag{9}$$

When calculating this integral using integration by parts, we obtain:

$$\int x^n \cdot e^{\alpha x} dx = e^{\alpha x} \cdot \sum_{k=0}^n \frac{(-1)^k x^{n-k} n!}{\alpha^{k+1} (n-k)!} \tag{10}$$

Another parameter is associated with terms - its complexity. It is complexity, together with the potential (individual) abilities of the learner, that affects the nature of the functions of learning to forget. In general, the term complexity is calculated through the depth of the hierarchy of the semantic network of the term definition.

Making the change of variables in the last expression, we arrive at the following form of representing the forgetting function:

$$Z(t | n, \alpha) = e^{-\alpha t} \cdot \sum_{k=0}^n \frac{\alpha t^k}{k!}, \tag{11}$$

which makes it possible by varying the parameters m and α to simulate various delays in forgetting and inflection points (see Fig. 6.b). To construct the learning function in the work, it is supposed to use the functions, the inverse forgetting functions $U(t)=1-Z(t)$ (see Fig. 6.a).

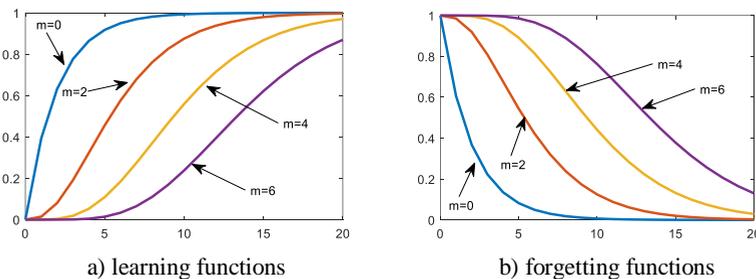


Fig. 6. Parameterization of the functions of learning and forgetting terms.

It should be noted that for the same term the parameters of the learning function $U(t/n, \alpha)$ and the function $Z(t/m, \beta)$ of forgetting can differ.

5 Recursive scheme for constructing the functions of learning forgetting

Based on the analysis of learning functions $U(t)$ and forgetting functions $Z(t)$ the paper proposes a piecewise functional representation of the learning-forgetting function $UZ(t/n, \alpha, m, \beta)$, i.e. in the process of studying modules that use given terms at a certain time interval, an increasing learning function is realized, and in the absence of such ones, a decreasing forgetting function $Z(t)$ is realized at the corresponding interval (see Fig. 7).

Of interest is the entire set of values of the learning-forgetting function of all terms of the curriculum. Therefore, to identify the function for the term, we introduce the indexing $UZ_{Term}(t/n, \alpha, m, \beta)$.

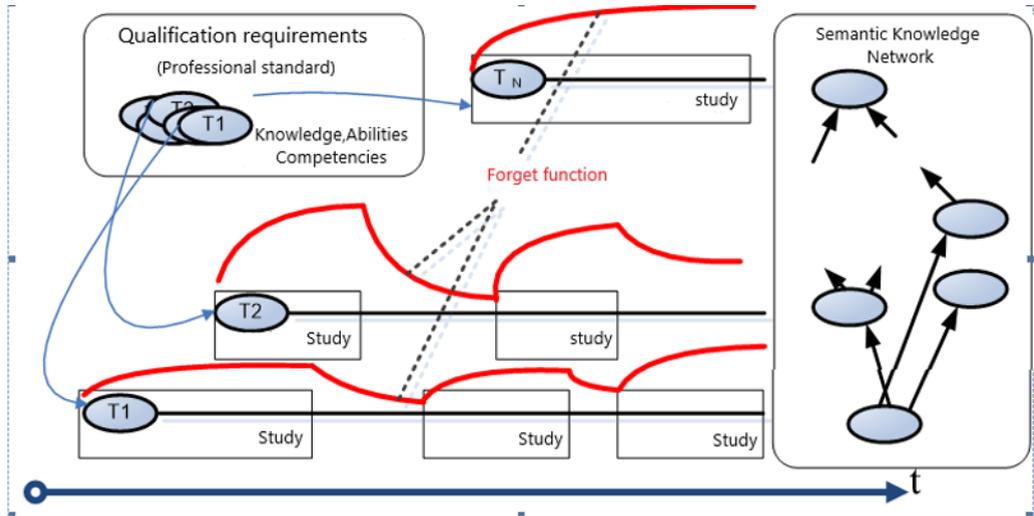


Fig. 7. The construction of learning-forgetting functions $UZ(t)$ terms of the curriculum.

When implementing the algorithmic scheme for generating the learning-forgetting function $UZ_{Term}(t/n, \alpha, m, \beta)$ the base sequence is the sequence U_{p_i} of logical zeros and ones corresponding to the absence or presence of modules with this term.

Each time when changing zero by one and vice versa at the time t_f , the reference point changes and when playing the corresponding function, the final value is saved. In this case, the continuation of the learning-forgetting function $UZ_{Term}(t/n, \alpha, m, \beta)$ will be realized through scaling of the initial learning functions:

$$UZ(t) = UZ(t_\phi) + (1 - UZ(t_\phi)) * U(t - t_\phi) \quad (12)$$

and forgetting:

$$UZ(t) = UZ(t_\phi) + (1 - UZ(t_\phi)) * Z(t - t_\phi). \quad (13)$$

When building a model of learning-forgetting functions, the start times for studying modules are taken from the calculated network model of the curriculum.

6 Variation in the sequence of reading modules of the educational program

Consider a sufficient simple network plan (see Fig. 8.a) of nine modules with three terms for which the learning and forgetting functions correspond to different complexity of the terms (see Fig. 8.b). For the first term (the most difficult) $\alpha=0.05$, $\beta=0.025$. For the second term $\alpha=0.1$, $\beta=0.02$. For the third term (the lightest) $\alpha=0.2$, $\beta=0.01$.

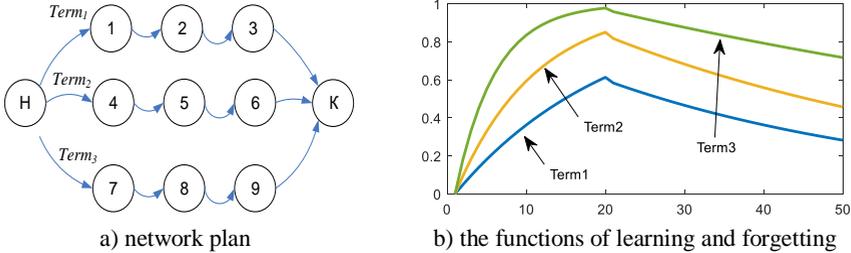


Fig. 8. Learning-forgetting functions for different sequence of reading modules.

For various permutations of the modules, we find the values of the learning-forgetting functions at the time the training is completed. Term 1 is used in modules 1, 2 and 3. Term 2 is used in modules 4, 5 and 6. Term 3 is used in modules 7, 8 and 9. If modules 1,2 and 3 with the most complex term are read at the very beginning, then modules 4, 5 and 6, and after that 7, 8 and 9, then upon completion of training the complex term is almost forgotten (see Fig. 9.a).

If the sequence of reading the modules is as follows, i.e. module 3 is read first, second - 5 and so on 9, 2, 6, 8, 1, 4, and 7, then the meaning of learning-forgetting functions will have a completely different look (see Fig. 9.b). Moreover, the values of the learning function of forgetting will be very close to each other and equal respectively to $UZ_1(t)=0.69$, $UZ_2(t)=0.62$, $UZ_3(t)=0.66$.

As a result, it was shown that the sequence of reading modules quite strongly affects the assimilation of various terms, which makes the solution to the problem of optimizing the choice of a sequence of modules, which is generated on the basis of the resource plan of the network plan, relevant.

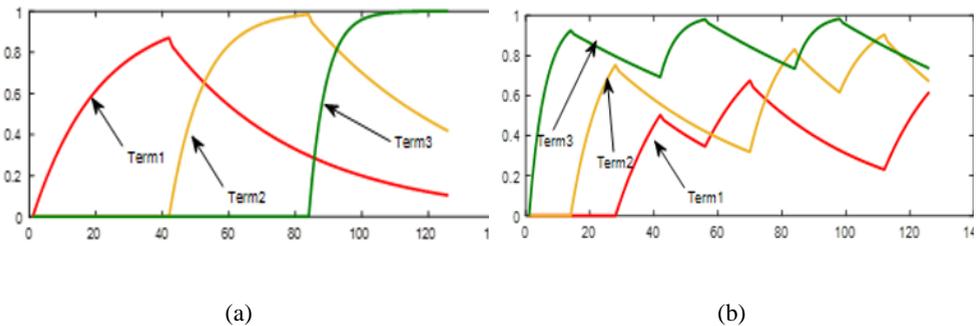


Fig. 9 Learning-forgetting functions for different sequence of reading modules

7 Conclusion

Thus, on the basis of the constructed models, each output term introduced in a module for each moment of time is associated with a numerical value that determines the degree of understanding of the term. Based on the obtained models of the forgetting function of each term, the paper proposes a multi-criteria statement of the problem of assessing the effectiveness of the training, which is as follows:

$$\forall i: UZ_i(T) \rightarrow \max, \quad (14)$$

where $UZ_i(T)$ – is the value of the learning-forgetting function of the i -th term at the time of completion (T) of the study of all modules. This is a classic multi-criteria task.

The basis for the formation of an integral criterion is the convolution of all functions by groups of classifying attributes of a module belonging to a certain direction. Each direction is assigned weight coefficients, which are transferred to all terms of the direction.

The optimization technique also includes a number of well-known methods for solving multicriteria problems (the ideal point method, the method of consecutive concessions, etc.). thus, by defining a performance criterion, it becomes possible to compare curricula, and therefore the possibility of their optimization.

References

1. G.G. Agadzhanov, V.S. Najdenko & V.S. Toiskin, *Nekotorye kriterii optimizacii uchebnyh planov*, Nauchnyj elektronnyj arhiv. URL: <http://econf.rae.ru/article/7755>.
2. V.Yu. Stroganov, O.B. Rogova, L.V. Ivanova, G.G. YAgudaev, *Formalizovannoe opisanie uchebnoho plana*, V mire nauchnyh otkrytij, Krasnoyarsk: NIC, No. 9 (21), pp. 16-28 (2011)
3. E.Yu. ZHazha, A.B. Nikolaev, D.V. Stroganov, E.Yu. Treshchetkina, L.V. Prihod'ko, *Formalizovannaya model' uchebnoho plana v zadache optimizacii individual'noj obrazovatel'noj traektorii*, Nauka i obrazovanie, No. 11, noyabr', pp. 351-361 (2012)
4. R.M. Harden, *AMEE Guide No. 21: Curriculum mapping: a tool for transparent and authentic teaching and learning*, Medical teacher, 23(2), pp. 123-137 (2001)
5. D. Spracklin-Reid & A. Fisher, *Curriculum Mapping In Engineering Education: Linking Attributes, Outcomes And Assessments*, Canadian Engineering Education Association (CEEA) (2014)
6. J. Seering, K. Willcox & L. Huang, *Mapping outcomes in an undergraduate aerospace engineering program*, American Society for Engineering Education (2015)
7. F. Rawle, T. Bowen, B. Murck & R. Hong, *Curriculum mapping across the disciplines: differences, approaches, and strategies*, Collected Essays on Learning and Teaching, 10, pp. 75-88 (2017)
8. K.E. Willcox & L. Huang, *Mapping the CDIO curriculum with network models, Preprint CDIO*, June 2017, 13th International CDIO Conference (2017)
9. D.A. Novikov, *Zakonomernosti iterativnogo naucheniya [Patterns of iterative learning]*, Moscow, Institute of management problems of the Russian Academy of Sciences, p. 77(1998)
10. L.L. Thurstone, *The learning curve equation*, Psychol. Monogr, **Vol. 26**, No. 3 (1919)

11. D.A. Novikov, *Setevie medeli v upravlenii: sbornik statey [Network models in management: collection of articles]*, of the Russian Academy of Sciences, V. A. Trapeznikov Institute of Probl. UPR. RAS, Moscow, Agues, p. 442: illustration. 21 (2011)
12. V.I. Voropaev, Ya.D. Gelrud, *Cyclic Alternative Network Models for Project Management*, 17th World Congress on Project Management, June 4–6 (2003)
13. E.R. Kasimov, E.V. Kuznetsova, *Event-marketing of higher education*, Business. Education. Law, Bulletin of the Volgograd Business Institute, No. 4 (33), pp. 345–351 (2015)
14. A.L. Gudkov, V.I. Ustich, A.A. Nedostup, *Development of training plans for basic educational programs of FSES of HVE*, Materials of the XVIII International scientific-methodical conference on February 17–18, 2011 «High intellectual technologies and innovations in scientific research universities», **Vol. 1**, SPb., Publishing house of Polytechnic University, pp. 23–24 (2011)
15. A.V. Gavrilenko, S.V. Rozova, T. A. Simonyan, *Selection of the mechanism for implementing improvement of the education quality control at higher schools*, Business. Education. Law, Bulletin of the Volgograd Business Institute, No. 2 (31), pp. 332–334 (2015)
16. N.A. Selezneva, *Quality of higher education as an object of systematic study*, Moscow, Research Center of the specialists training quality issues, p. 80 (2001)
17. A.A. Ryzhkin, O.A. Polushkin, *Process of specialists training at higher school as a complex system*, Materials of the X International Conference «Modern technologies of training «SRT-2004», **Vol. 1**, SPb., LETI, pp. 24–26 (2004)
18. O. Polushkin, G. Kanygin, M. Polushkina, *Quantitative evaluation of quality and modeling of the academic discipline preparation process*, VII International Scientific practical conference «Naynovite evropeyskata comprehension on science», **Vol. 18**, Pedagogical Sciences, Sofia: «Byal GRAD BG» LTD, pp. 6–12 (2011)