УДК 004.75

K. P. STORCHAK[1];
V. R. MYKOLAICHUK[2]
[1] State University of Telecommunications, Kyiv
[2] National Aviation University, Kyiv

ADAPTIVE COMPUTER TESTING BASED ON THE GRADIENT METHODS

The article deals with the gradient methods in adaptive computer testing for assessing the level of knowledge in the course of distant learning. It describes the possibility of increasing the accuracy of evaluation by determining the level of complexity of test questions in accordance with the current determined level of knowledge as well as ensuring the motivation of the test takers. The methodological approach to adaptive computer testing is described and the corresponding algorithm is suggested. The results of evaluation of the expected increase of testing efficiency by using adaptive computer testing on the example of development of a training course for the system of distant learning are analyzed.

Keywords: adaptive computer testing; distant learning; assessment of the acquired knowledge level; gradient methods; effectiveness of the testing algorithm.

Problem statement

Integration of Ukraine into the European community actualises the issue of improving the quality of specialists' training, identifies the search for the new forms and methods for organizing the educational process, application of the advanced teaching technologies, and the creation of means for effective evaluation of the level of knowledge acquired during the training courses by using a computer testing. To solve this problem by using traditional approaches [1] it is necessary to increase the time for testing and use a large set of test questions of varying degrees of complexity. On the other hand, the need to increase the time for direct study of educational material and prevention of users' motivation decrease when they answer too simple or too complex test questions requires different approaches.

The solution to the above contradiction seeks for the use of new approaches to computer testing. In particular, the use of adaptive testing, especially as a kind of testing in which the order of questions and their complexity depend on the answers of the test taker (further tested) on the previous questions, looks promising [2].

Analysis of the recent research and publications

The basic theoretical positions of adaptive testing are discussed in Adaptive Tests: General Provisions [2]. The definition of adaptive testing is given, a comparison of traditional and adaptive tests is conducted, and general approaches to adaptive testing are presented. Later in Using Adaptive Tests in Intelligent Control Systems Knowledge [3], based on the known single- and multiparametric models, approaches to streamline test questions in terms of complexity and proposed adaptive testing algorithms for use in distant learning are discussed. From the results of the analysis of these publications, it follows to the conclusion about the possibility of increasing the adaptability of testing with the accuracy of the assessment, as well as ensuring the motivation of the test takers tested by setting the level of complexity of test questions in accordance with the current level of knowledge.

At the same time, existing publications do not address the issue of the choice of the complexity level of each subsequent question during adaptive testing, according to the previous responses of the test takers. The need for solving these issues is acutely encountered during the practical implementation of adaptive testing in distant learning courses.

Purpose of the article

Given the iterative character of the testing process and the discreteness of the evaluation, for the solution of the above questions, it seems appropriate to use the gradient approach [4], which essence is to determine the next approximation to a minimum of some functional dependence f from the previous in the direction opposite to its gradient ∇f.

Taking into account the above, the purpose of the article is to highlight the gradient approach for adaptive computer testing in assessing knowledge in distant learning.

Presentation of the main research material

Let the set of test questions be defined (we will assume that all questions meet the requirements of validity and reliability [1]). Based on predefined levels of complexity of questions (for example, according to [3]), we introduce the equivalence relation H on Q, which ensures the classification of test questions according to their complexity: $H: Q \bigcap Q/H$ where Q/H is the corresponding ordered quotient set whose elements are classes of equivalence (complexity issues) of the relation H, which we denote by h_i, $i \in \overline{1,|H|}$.

On the basis of the ordering of the set H, we define the corresponding metric $\mu(h_j, h_k) = h_j, h_k = \left|\{h\}: h \prec h_j\right|\left|\{h\}: h \prec h_k\right|$. Let $h_i(Q) = \{q \in Q: H(q) = h_i\}$. We introduce additional parameters of difficulty levels h_i: the number of correct answers to fulfil the passing condition of i-th level h_i^r, the number of incorrect answers to fulfil the level «failure» h_i^W (the corresponding current numbers of correct and incorrect answers at the level will be marked as q_i^r q_i^W).

Also, we will denote $a_n, n \in \overline{1,|Q|}$ as the answer to the current n-th test question, and the set of answers to the first n questions $\{a_n\} = \{a_k\}: k \in \overline{1,n}$. Suppose the presence of someone who tests a certain level of knowledge h_u, $h_u \in H$.

Then the task of choosing the next test question can be formulated as minimizing the differences in the complexity of each subsequent issue and the level of knowledge of the tested, based on its previous answers:

$$q_{n+1} \in h_i(Q): \min_{\{a_n\}}(\mu(h_u, h_i)). \tag{1}$$

The peculiarity of solving the problem (1) lies in the uncertainty of the value h_u by the time the test is completed.

To solve the problem (1) we will give a test in the form of a certain functional dependence t, which determines the level of knowledge of the person who is tested (based on the set of answers A):

$$h_t = t(A). \tag{2}$$

The analysis of the possible dependence properties (2) in accordance with the classical and adaptive approach to testing allows us to draw the following conclusions:

- for classical testing, the following dependencies are executed:

$$|A| = |Q| \ge \sum_i \left(h_i^r + h_i^W - 1\right), \;\; h_t = \max_i h_i: \; q_i^r \ge h_i^r \left| \exists h_k: \; q_k^W \ge h_k^W, k \in \overline{1,i}\right.,$$

i. e. the test is conducted on the whole set of questions, and the level of knowledge is determined according to the highest $\left(\max_i h_i\right)$ passed $\left(q_i^r \ge h_i^r\right)$ level of difficulty h_i, provided that there is no precedent for the level failure $\left(\forall h_k: \; q_k^W < h_k^W, \; k \in \overline{1,i}\right)$;

- the presence of non-structural inequalities in the criteria of passing/failure of the level indicates the possibility of reducing the number of questions asked during the computer testing to the limit $|A| \le \left|\sum_i \left(h_i^r + h_i^W - 1\right)\right|$, through the passage of other issues level to achieve the condition, $q_i^r = h_i^r$ as well as termination of testing in the case $q_i^W = h_i^W$ (this approach will be called elementary adapted testing);
- the greatest interest in terms of solving the problem (1) acquires the testing process in the pre-criterion interval $\left(q_i^r < h_i^r, \; q_i^W < h_i^W\right)$, which will be considered below.

The essence of the proposed approach is to initiate an early transition to the next/previous level of questions based on the assessment of current responses in terms of providing the required «rate» of passage/failure of the level.

For this purpose, we will submit the testing process at one difficulty level in the form of a graphical mapping on the decimal line [−1; +1] (fig. 1):

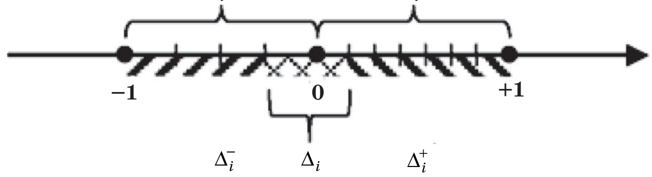

Fig. 1. Graphical representation of the testing process at the i-th level of difficulty

In fig. 1 point 0 corresponds to the initial value $t(A)$, +1 and −1 the passage and failure of the level respectively. In this case, the intervals [−1; 0] and [0; +1] are divided by the intervals of the corresponding length $\frac{1}{h_i^W}, \frac{1}{h_i^r}$. Also, the selected interval of the let-down $\Delta_i = \left[0 - \frac{1}{h_i^W}, 0 + \frac{1}{h_i^r}\right]$, as well as the corresponding intervals Δ_i^-, Δ_i^+ will determine the conditions of abandonment at the current level or transition to another level during the adaptive testing. The specified intervals can be the basis for the solution of the problem (1):

$$q_{n+1} \in \begin{cases} h_{i+1} \leftarrow t(A) \in \Delta_i^+, \\ h_i \leftarrow t(A) \in \Delta_i, \\ h_{i-1} \leftarrow t(A) \in \Delta_i^-. \end{cases} \quad (3)$$

At the same time, in the expression (3) the «interval» approach to solving the problem (1) does not take into account the ratio of the acceptable number of correct and incorrect answers, and therefore needs to be improved.

In order to improve the approach (3), the application of gradient methods is proposed. Note that after each correct answer there is a change in the value of the expression $t(A)$ in $\frac{1}{h_i^r}$ at positive direction, after the wrong one, in $\frac{1}{h_i^W}$ at negative direction (see fig. 1), which makes it possible to determine the functional dependence $t(A)$ in the following way:

$$t(A) = \begin{cases} h_{i-1} \leftarrow q_i^W = h_i^W, \\ h_i + \dfrac{q_i^r}{h_i^r} - \dfrac{q_i^W}{h_i^W} \leftarrow \left(q_i^W < h_i^W\right) \cup \left(q_i^r < h_i^r\right), \\ h_{i+1} \leftarrow q_i^r = h_i^r. \end{cases} \quad (4)$$

According to expression (4), it is possible to determine the direction and magnitude of the increase for functional dependence $t(A)$. This allows entering $t(A)$ for determination the gradient:

$$\nabla t(A) = \frac{\dfrac{\Delta q_i^r}{h_i^r} - \dfrac{\Delta q_i^W}{h_i^W}}{\Delta |A|} j, \quad (5)$$

where j is the unit vector of the transition to the next level with the beginning at point 0 and the vertex at the point +1 in accordance with fig. 1.

According to Expression (5) it is possible to determine the direction of the testing process (evaluation) from the point of view of determining the level of knowledge of the test based on the difference in the ratio of correct and incorrect answers. The positive direction of the gradient indicates the growth of a certain level of knowledge, the negative — to decrease. This makes it possible to change the current level of knowledge until the boundary conditions are reached (4). To do this, define the gradient $\nabla_i^{\min}$ as the minimum required direction of growth $t(A)$ for successful passing of some i-th level:

$$\nabla_i^{\min} = \frac{1 - \dfrac{h_i^W - 1}{h_i^W}}{h_i^r + h_i^W - 1} j = \frac{i}{h_i^W \left(h_i^r + h_i^W - 1\right)}. \quad (6)$$

By analogy with the approach (3), we define the boundary values of the gradient, which will determine the conditions for the transition to another level, by changing the expression (6) by adding one correct/incorrect to the set of answers:

$$\nabla_i^+ = \frac{\dfrac{h_i^r + 1}{h_i^r} - \dfrac{h_i^W - 1}{h_i^W}}{h_i^r + h_i^W} j = \frac{1}{h_i^r \cdot h_i^W} j,$$
$$\nabla_i^- = \frac{\dfrac{h_i^r}{h_i^r} - \dfrac{h_i^W}{h_i^W}}{h_i^r + h_i^W} j = 0. \quad (7)$$

The above expressions (4)–(7) provide the possibility of applying tested gradient methods to solve the problem (1):

$$q_{n+1} \in \begin{cases} h_{i+1} \leftarrow \nabla t(A) > \nabla_i^+, \\ h_i \leftarrow \nabla_i^- \le \nabla t(A) \le \nabla_i^+, \\ h_{i-1} \leftarrow \nabla t(A) < \nabla_i^-. \end{cases} \quad (8)$$

The proposed approach (4)–(8) allows you to choose the next question, taking into account the previous answers of the person who is tested, that is, provides a solution to the problem (1).

For software implementation of adaptive testing, LMS Moodle and Adobe Captivate software were used as the main software tools used in the distant learning system (fig. 2).

This figure shows the auxiliary (hidden) slide that displays the results of the test as a whole (variable scores — $$ MainScore $$ and achieved level — $$ Rate $$), as well as by levels (status of *i*-th level [failed, passed, missed according to the adaptive algorithm] — $$ LiStatus8 $$, the number of correct answers — $$ LiPoint $$, and the number of incorrect — $$ LiMist $$).

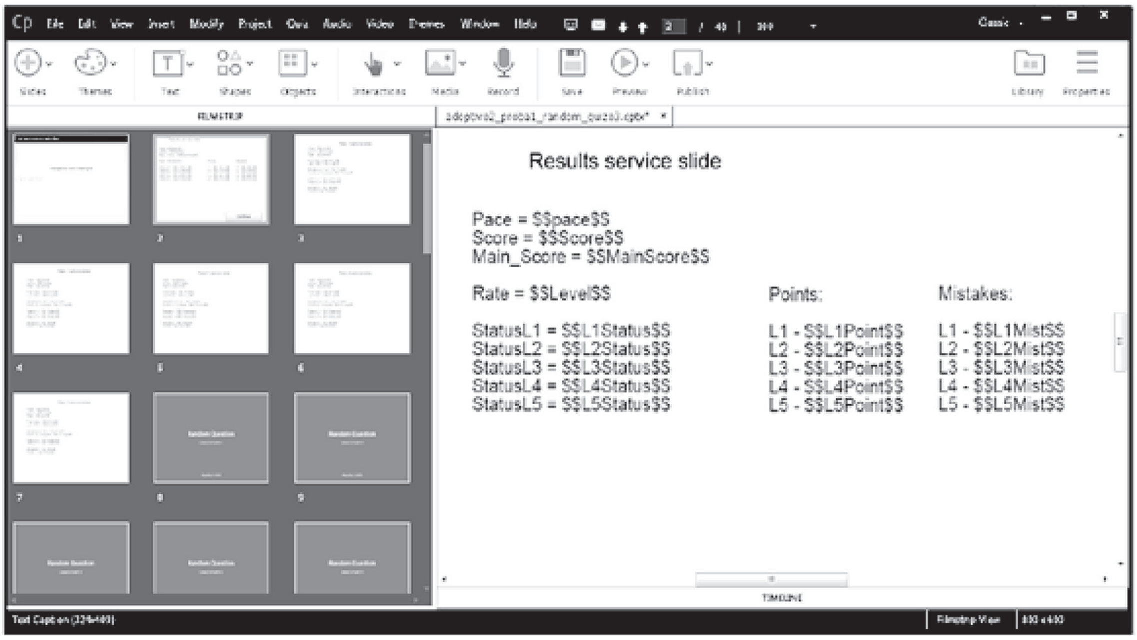

Fig. 2. Creating software for adaptive testing in an integrated development environment Adobe Captivate

Similar auxiliary slides are created for each of the five predicted levels (the number of levels selected according to the testing requirements for STANAG6001 [1, 1+, 2, 2+, 3]). The display of auxiliary slides while debugging the program allows you to track the performance of an adaptive algorithm and pre-evaluate its effectiveness.

For the purpose of comparative analysis of the proposed approach in comparison with the classical, elementary adaptive and «interval» approaches, statistical simulation was carried out using the Monte Carlo method.

For statistical simulation a test with the following characteristics are selected $|H| = 5$, $\forall h_i$: $|h_i(Q)| = 10$. The statistical degree of preparation of the «tested» according to the existing criteria for evaluation of the tests was determined by the probability pi of the correct answer to the question of the 1st level received: 0,9 — with the degree of preparation that exceeds; 0,7 — satisfies; 0,5 — partially satisfying; 0,25 — does not meet the criteria of the current level (the possibility of «guessing» the correct answer taking into account the availability of four options for response) is taken into account. Accordingly, it is possible to enter the probabilistic profile of the test $P = \{p_i\}$. In particular, for the simulation, $P = \{0{,}9; 0{,}8; 0{,}7; 0{,}5; 0{,}25\}$ was taken with the expected successful achievement of the third level by the results of testing.

Subsequently, 10,000 implementations of the testing process using the MatCad software environment were conducted. During each implementation, using the random number sensor, the results of the answers to the test questions were determined and the test results were determined for each of the above approaches.

The main results of statistical simulation are given in the table.

Results of statistical simulation

Approach name	Number of questions in the test	Average number of questions to determine the result	Correlation coefficient of test results
Classic	50	50	1
Elementary Adaptive	50	42	1
Interval adaptive	50	33	0,99
Gradient Adaptive	50	25	0,98

According to the results obtained, the proposed approach yields a significant time gain (by reducing the required number of questions for the outcome) compared to the classical (50%) and elementary adaptive (30%) approaches, while maintaining the accuracy of the assessment within 3Compared with the interval approach, the time reduction is set at 8%, which, for the selected confidence value of 0,95, allows us to conclude on the increase in the time value of the effectiveness of the evaluation algorithm using the gradient approach. The efficiency growth rate of the algorithm proposed in expressions (2)–(8) will increase with the increase in the number of assessed levels and test questions.

Conclusions and perspectives of further research

Thus, the preliminary assessment of the results of the use of the gradient approach to adaptive testing during the assessment of the quality of training in distant learning courses allows us to conclude that achieving a reduced test time with adherence to the adequacy of the assessment.

This will ensure the efficient use of educational resources on distant learning or increase the effectiveness of assessment by increasing the test. Further researches should be conducted in the direction of elaboration of an integrated evaluation algorithm in the presence of a plurality of the estimated parameters with their joint consideration in each of the test questions.

References

*1. **Ositrov V. V.** Terms and educational foundations control of troops of foreign language using modern technology of computer adapted testing [Umovy ta dydaktychni osnovy kontroliu rivni avolodinnia inozemnoiu movoiu viiskovosluzhbovtsiv z vykorystanniam suchasnoi tekhnolohii kompiuterno adaptovanoh otestuvannia] // Visnyk NAOU. 2009. No. 2(10). P. 56–62.*

*2. **Fedoruk P. I.** Adaptive tests: general provisions [Adaptyvni testy: zahalni polozhennia] // Mathematical Machines and Systems. 2008. No. 1. P. 115–127.*

*3. **Fedoruk P. I.** Using adaptive tests in intelligent control systems knowledge [Vykorystannia adaptyvnykh testiv v intelektualnykh systemakh kontroliu znan] // Shtuchnyi intelekt. 2008. No. 3. P. 380–387.*

*4. **STANAG 600.** [Electronic resource].— http://www.natobilc.org/files/file/6001EFed05.pdf.*

Рецензент: доктор техн. наук, професор **К. С. Козелкова**, Державний університет телекомунікацій, Київ.

К. П. Сторчак, В. Р. Миколайчук

АДАПТИВНЕ КОМП'ЮТЕРНЕ ТЕСТУВАННЯ НА ОСНОВІ ГРАДІЄНТНИХ МЕТОДІВ

У статті описано градієнтні методи адаптивного комп'ютерного тестування під час оцінювання знань у ході дистанційного навчання. Визначено можливість підвищення точності оцінювання та забезпечення збереження мотивації тих, хто тестується, за рахунок встановлення рівня складності тестових питань відповідно до поточного рівня знань. Запропоновано методологічний підхід до адаптивного комп'ютерного тестування та відповідний алгоритм. Наведено результати оцінювання очікуваного підвищення ефективності тестування за рахунок використання адаптивного комп'ютерного тестування на прикладі розробки навчального курсу для системи дистанційного навчання.

Ключові слова: адаптивне комп'ютерне тестування; дистанційне навчання; оцінювання рівня засвоєних знань; градієнтні методи; ефективність алгоритму тестування.

К. П. Сторчак, В. Р. Миколайчук

АДАПТИВНОЕ КОМПЬЮТЕРНОЕ ТЕСТИРОВАНИЕ НА ОСНОВЕ ГРАДИЕНТНЫХ МЕТОДОВ

В статье описаны градиентные методы адаптивного компьютерного тестирования при оценке знаний в ходе дистанционного обучения. Определена возможность повышения точности оценки и обеспечения сохранения мотивации тех, кто тестируется, за счет установки уровня сложности тестовых вопросов в соответствии с текущим уровнем знаний. Предложен методологический подход к адаптивному компьютерному тестированию и соответствующий алгоритм. Приведены результаты оценки ожидаемого повышения эффективности тестирования за счет использования адаптивного компьютерного тестирования на примере разработки учебного курса для системы дистанционного обучения.

Ключевые слова: дистанционное обучение; оценивание уровня усвоенных знаний; адаптивное компьютерное тестирование; градиентные методы; эффективность алгоритма тестирования.