Journal of Engineering Science and Technology
Vol. 14, No. 5 (2019) 2485 - 2498
© School of Engineering, Taylor's University

# ADAPTIVE TESTING MODEL APPROACH BASED ON BIRNBAUM MODEL AND MARKOV MODEL

HOAI PHUONG DANG*, VAN HA MAI

Information Technology Faculty, The University of Danang
- University of Science and Technology (UD-DUT), Viet Nam
*Corresponding Author: dhphuong@dut.udn.vn

**Abstract**

Computerized adaptive testing is widely used in the testing and assessment of the level of learners' competency. Popular computerized adaptive testing systems now use the mathematical models of Item Response Theory based on the relationship between the ability of examinees and item parameters. However, Item Response Theory does not take into account the impact between previous answers and the next item selection. The item parameters principally rely on probability methods of Classical Test Theory. This article proposes the combination of Item Response Theory (Birnbaum model) and Markov chain to calculate the dependency of answer set during the quiz process. Concurrently, using Hooke - Jeeves direct search method within the limited range of parameters to assess the set of item parameters.

Keywords: Computerized adaptive testing, Hooke-Jeeves method, Item parameters, Item response, Markov chain, Theory.

## 1. Introduction

An essential component of e-learning is knowledge control. Currently, various methods and algorithms for computer testing have been developed, yet most of them are based on Classical Test Theory (CTT) [1]. The most adequacy is adaptive testing [2], which involves changing the sequence, content, difficulty of items in the real testing process, depending on the actions of the examinee. While implementing adaptive testing, the order and numbers of items in a quiz will differ for different examinees based on their ability (strong, medium, weak level).

At present, Computerized Adaptive Testing (CAT) [3] usually uses the Item Response Theory (IRT) [4-6], of which, the next item selection depends on the evaluation of the examinee's ability calculated by the base of the previously received responses. To construct the test trajectory, the approach of maximization information is used where each item selection from the bank of items owns most information to ensure the accurate assessment of the examinee's ability. The process testing permits the item selection for the examinee's ability adaptively, which could reduce the quiz time as well as obtain the accuracy of assessments.

Nonetheless, item parameters of the IRT model are assessed primarily by probability methods of CTT [7, 8] leading to the outcome that its accuracy is not high. Moreover, this model does not examine the dependency of earlier answer collection during the quiz. Therefore, the authors offer an adaptive testing model based on the IRT and Markov chain to solve the above problems.

## 2. Related Work

According to Lord [9], IRT is based on the latent class analysis]. In the IRT, a logistic function is used to describes the dependency of the probability of a correct response to the item relied on the both of examinee's ability and the set of item parameters [10]. Mathematically, the theory allows the connection between the answers and the assigned items constructed by statistics of the items and the examinee's ability.

### 2.1. CAT based on Birnbaum model

With adaptive testing, the latent properties of the individual are understood as the examinee's ability, which represents the level of their preparation for the assigned items to them. Overall, in the IRT, the probability of a correct answer depends on the set of properties, which differs depending on the used mathematical models (Rasch model, two parameters logistic model - 2PL, three parameters logistic model - 3PL) [11]. According to the Birnbaum model [11], the probability of a correct answer to a given item $i^{th}$ relies on the examinee's ability and the three parameters: $a_i$ is an item discrimination parameter, $b_i$ is an item difficulty parameter and $c_i$ is an item pseudo-guessing parameter ($0 < c_i \leq 1$), which estimates the probability of a correct answer to item $i^{th}$ by chance.

In accordance with the model, the probability of a correct answer to item $i^{th}$ answered by the $j^{th}$ examinee can be expressed as follows [11]:

$$p_j\left(u_{ji} = 1 | \theta_j, a_i, b_i, c_i\right) = c_i + (1 - c_i)\frac{1}{1+e^{-1.7a_i\left(\theta_j - b_i\right)}} \quad (1)$$

where $u_i$ is an answer of the $j^{th}$ examinee to item $i^{th}$ ($u_{ji}$ = 1, if the answer is correct, $u_{ji} = 0$ otherwise) and $\theta_j$ - the ability of $j^{th}$ examinee.

From the known parameters ($a_i$, $b_i$, $c_i$) of item $i^{th}$ for the probability value Eq. (1), we can plot the probability response of the correct answer based on the examinee's ability $\theta$ (as shown in Fig. 1).

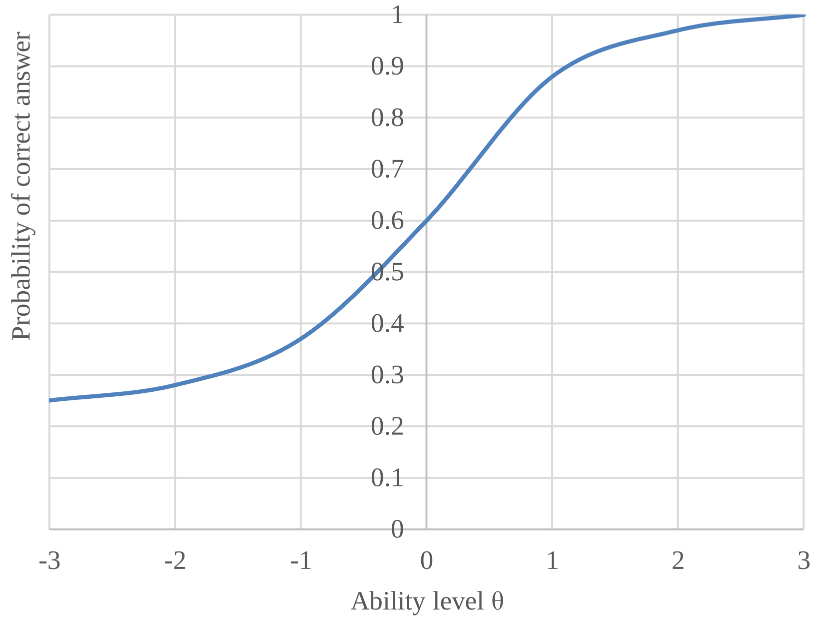


**Fig. 1. Graph of probability of a correct answer to an item $i^{th}$ based on examinee's ability $\theta$ for $a_i = 1$, $b_i = 0$, $c_i = 0.25$.**

The process of CAT in the Birnbaum model is iterative and has the following steps (Fig. 2):

- **Step 1**: Set the initial value of the examinee's ability.
- **Step** 2: Repeat component A (steps 3 through 5) until the stop criterion is reached. The stop criterion is the change rate of examinee's ability, which is smaller than the given value.
- **Step** 3: All items, which have not been assigned yet for the examinee, are evaluated to choose the best item for the current examinee's ability.
- **Step 4**: The item selection is allocated for the examinee to answer.
- **Step 5**: A new assessment of the examinee's ability is calculated based on the received response.
- **Step 6**: The evaluation of new achieved capacity is used to evaluate the examinee's ability.

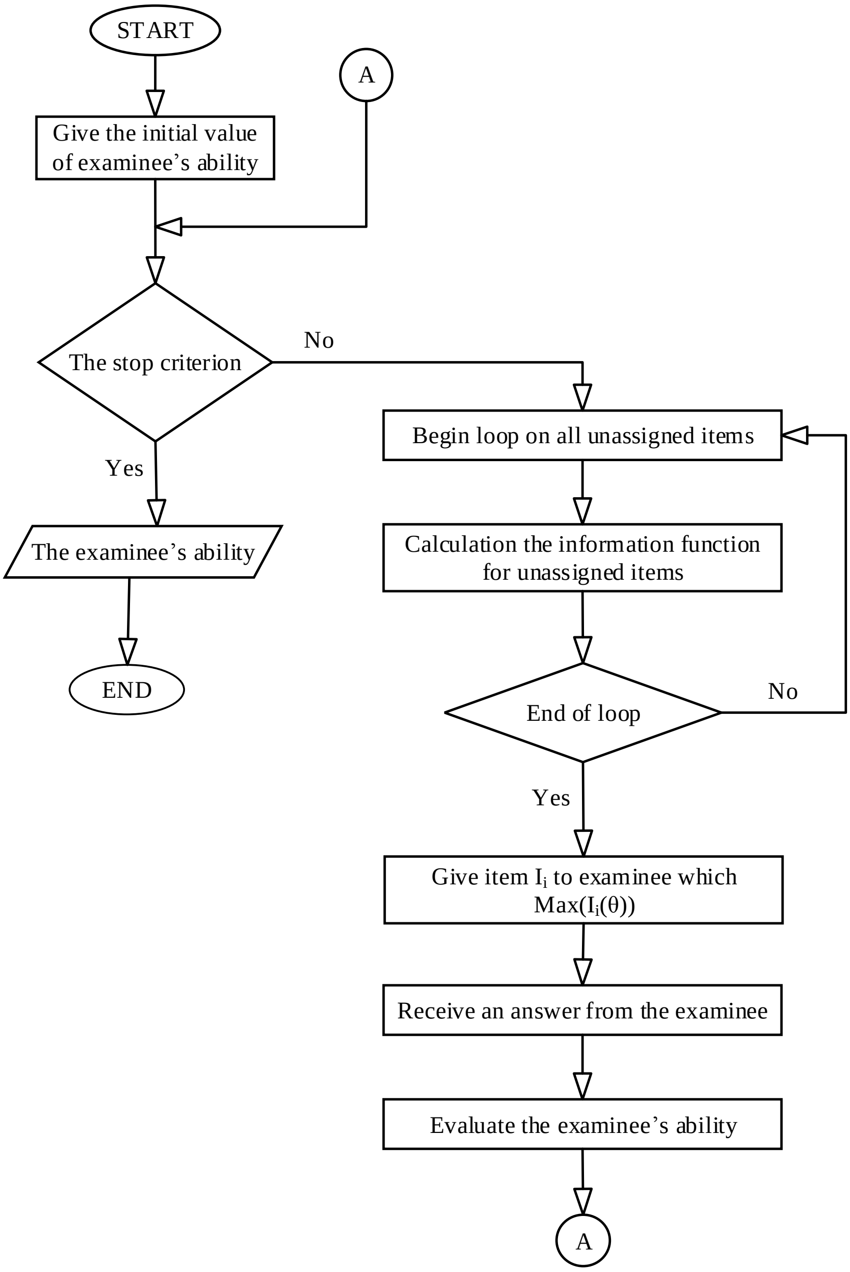


**Fig. 2. Algorithm adaptive testing of Birnbaum model.**

The first step is to assess the examinee's ability either from the results of previous tests or to assume it as an average value (it usually equal to 0). The second step is to select the best item determined by the given criteria. It is because an item will provide little information about the evaluation of ability if it is too complicated or too easy for the examinee. The items should have the most information to assess the examinee's ability. To deal with this problem, in the Birnbaum model, depending on the evaluation of the examinee's ability, the information function [12] value of item $i^{th}$ is computed by the formula:

$$I_i(\theta_j) = \frac{p_i'(u^{(i)}=1|\theta_j,a_i,b_i,c_i)^2}{p_i(u^{(i)}=1|\theta_j,a_i,b_i,c_i)\left(1-p_i(u^{(i)}=1|\theta_j,a_i,b_i,c_i)\right)'} \quad (2)$$

where: $p_i(u^{(i)} = 1|\theta_j, a_i, b_i, c_i)$ is a probability of a correct answer to item $i^{th}$ of $j^{th}$ examinee $(u^{(i)} = 1)$, $p_i{}'\ (u^{(i)} = 1|\theta_j, a_i, b_i, c_i)$ is the first derivative of the probability function $p_i(u^{(i)} = 1|\theta_j, a_i, b_i, c_i)$ and $(a_i, b_i, c_i)$ - set parameters of item $i^{th}$.

Thus, in step 1 of the adaptive testing algorithm based on the Birnbaum model, the information function $I_i(\theta_j)$ can be calculated for each of the ungiven items at the ability of $j^{th}$ examinee $\theta_j$. The most common approach to item selection is to find the one with the maximum value of this function.

In step 3, a new evaluation of the ability of $j^{th}$ examinee $\theta_j$ with the maximum likelihood method [13] to assess this ability is made, i.e., to find such an assessment of $\theta_j$ where the likelihood function value will be maximal.

$$\left(u^{(1)}, u^{(2)}, \dots, u^{(n)}, \theta_j\right) = \prod_{k=1}^{n} p\left(u^{(k)}|\theta_j, a_k, b_k, c_k\right) \rightarrow Max \quad (3)$$

where $u^{(1)}$, $u^{(2)}$, …, $u^{(n)}$ is a set answer of an $j^{th}$ examinee (get value 1 if answer correct and 0 with otherwise), and $(a_k, b_k, c_k)$ is the parameter set of the item, which assigned to the examinee.

The above method uses a probabilistic model based on the statistical results of data processing in the quiz bank, which makes it possible to acquire an objective assessment of the examinee's ability. However, there is a disadvantage in the model since it does not take into account the possible influence of the set of previous answers and next item selection.

### 2.2. Adaptive testing model (the Birnbaum model and Markov chain combination)

As mentioned, the Birnbaum model can be significantly supplemented by correlation analysis of the examinee's answers. In this paper, we propose a model adaptive testing in the form of a homogeneous elementary Markov chain [14] for a binary sequence of responses.

While applying this mathematical apparatus, it is offered to enter two states: $u_i^{(k)} = 1$ when the examinee on the item $k^{th}$ answers correctly, and $u_i^{(k)} = 0$ - the state when it answers incorrectly, the transition of the system from one state to another can be described by a homogeneous Markov chain defined by the matrix of transition probabilities.

$$P = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} \quad (4)$$

where $p_{00}$ is the probability of a transition from an incorrect response to an incorrect response, $p_{01}$ is the probability of a transition from an incorrect answer to a correct one, $p_{10}$ is the probability of transition from the correct answer to the wrong one, and $p_{11}$ is the probability of transition from the correct answer to the correct one.

Figure 3 shows the state transition diagram for the above Markov chain. In this diagram, there are two possible states $u_i^{(k)} = 1$ and $u_i^{(k)} = 0$, and the arrows from each state to other states show the transition probabilities $p_{ij}$. When there is no arrow from state $i$ to state $j$, it means that $p_{ij} = 0$.

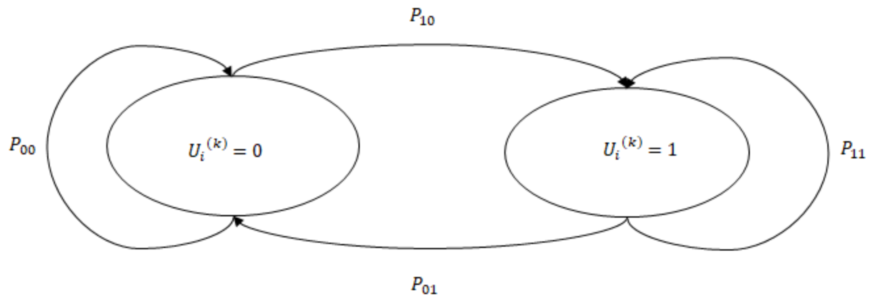


**Fig. 3. A state transition diagram.**

In general, the model analyses the sequence of responses must take into account the dependence of the transition matrix on the item selection of examinee at each step. In this paper, we propose a simpler model, which use the Kolmogorov-Chapman equation [15], one can find the probability ability to go from one state to another for any number steps $n$:

$$P(n) = P^n \tag{5}$$

It is clear that the transition probability is unchanged for different probability values at each step with the condition of item number sufficiently. Furthermore, for the matrix of transition probabilities, it is necessary to specify $p(u^{(0)} = 1)$ the initial probability distribution of the correct answers at the step $k = 0$.

The Birnbaum model changes according to the Markov dependence [15, 16], where the probability of a correct answer to item $i^{th}$ of $j^{th}$ examinee will be:

$$p_{ji}\left(u^{(i)} = 1 | u^{(i-1)} = 0, \theta_j\right) = c_i + (1 - c_i)\frac{1}{1+e^{-1.7a0_i\left(\theta_j - b0_i\right)}} \tag{6}$$

$$p_{ji}\left(u^{(i)} = 1 | u^{(i-1)} = 1, \theta_j\right) = c_i + (1 - c_i)\frac{1}{1+e^{-1.7a1_i\left(\theta_j - b1_i\right)}} \tag{7}$$

where $p_{ji}\left(u^{(i)} = 1 | u^{(i-1)} = 0, \theta_j\right)$ is a probability of a correct answer to item $i^{th}$ to $j^{th}$ examinee if an $j^{th}$ examinee submits incorrect answer for item $(i - 1)^{th}$. Otherwise, $p_{ji}\left(u^{(i)} = 1 | u^{(i-1)} = 0, \theta_j\right)$ is a probability of a correct answer to item $i^{th}$ to $j^{th}$ examinee if an $j^{th}$ examinee submits correct answer for item $(i - 1)^{th}$, and $(a0_i, a1_i, b0_i, b1_i, c_i)$ is the parameter set of item $i^{th}$ in both cases.

Similarly, the procedure for calculating the value of the information function [16] also changes:

$$I_i\left(\theta_j | u^{(i-1)} = 0\right) = \frac{{p_i}'\left(u^{(i)}=1|u^{(i-1)}=0,\theta_j\right)^2}{p_i\left(u^{(i)}=1|u^{(i-1)}=0,\theta_j\right)\left(1-p_i\left(u^{(i)}=1|u^{(i-1)}=0,\theta_j\right)\right)'} \tag{8}$$

and:

$$I_i\left(\theta_j | u^{(i-1)} = 1\right) = \frac{{p_i}'\left(u^{(i)}=1|u^{(i-1)}=1,\theta_j\right)^2}{p_i\left(u^{(i)}=1|u^{(k-1)}=1,\theta_j\right)\left(1-p_i\left(u^{(i)}=1|u^{(i-1)}=1,\theta_j\right)\right)'} \tag{9}$$

if the answer was correct.

The Markov chain is still in use for calculating the maximum likelihood function for estimating the current $j^{th}$ examinee's ability with the Birnbaum model. Thus, if the response vector of $j^{th}$ examinee $U = \{u^{(1)}, u^{(2)}, ..., u^{(n)}\}$ is obtained for testing, then in the case of the Birnbaum model the maximum likelihood function is calculated according to Eqs. (3), (6) and (7):

$$f\left(u^{(1)}, u^{(2)}, \dots, u^{(n)}, \theta_j\right) = p\left(u^{(0)} = 1|\theta_j\right).p\left(u^{(1)} = 1|u^{(0)}, \theta_j\right).p\left(u^{(2)} = 1|u^{(1)}, \theta_j\right) \dots p\left(u^{(n)} = 1|u^{(n-1)}, \theta_j\right) \to Max \quad (10)$$

The use of a homogeneous Markov chain is a generalization of the Birnbaum model. With the help of the described model, additional characteristics of the adaptive testing process can be introduced, such as the probability of the first return of the testing process to a given state (initial state) through *k* steps, the average number of steps of the first return to the specified state, and others.

We suggest the application of Markov chain for interval estimation of the examinee's ability under the Birnbaum model. Thus, there are four stages in the model, two of them are absorbing, namely, state U is "unsatisfactory" and state E is "excellent". State *S* shows that the level of the examinee is "satisfactory", and state *G*, that the level is "good". The matrix of transition probabilities is as follows:

$$P = \begin{matrix} \\ U \\ S \\ G \\ E \end{matrix} \begin{matrix} \begin{matrix} U & S & G & E \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ p_{SU} & p_{SS} & p_{SG} & 0 \\ 0 & p_{GS} & p_{GG} & p_{GE} \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix} \quad (11)$$

the peculiarity of this matrix is that the *S* cannot get into *E*, but *G* in *U*.

## 2.3. Estimation of item parameters

Item parameters for an adaptive testing model can be obtained by the statistical method of CTT, i.e. when information is received about the answers of an examinee to all items of the bank. The guessing parameter ci of the Birnbaum model and the guessing parameters *ci* for the proposed model are given by constants (normally equal to 0.25). Accordingly, these constants present the probability of guessing a correct answer, which is $\frac{m}{n}$, where *m* is the number of correct answers and *n* is the number of answers of assigned items.

In order to evaluate item parameters in this paper, it is proposed to use the method of least squares. In this case, each examinee, who passes the quiz, falls into one of the *M* groups determined by the range of the examinee's ability. Regarding each item $i^{th}$ of the item bank and the average ability of the examinee's group $j^{th}$, the estimation of the set of item parameters *i* are found by minimizing the following function:

$$g(a_i, b_i, c_i) = \sum_{j=1}^{M}\left(p\left(u = 1|\widehat{\theta_j}, a_i, b_i, c_i\right) - \widehat{p_{ij}}\right)^2 \to min \quad (12)$$

where *M* is the number of examinee groups participates in quiz would be divided based on the results of the examinee's ability, ($a_i$, $b_i$) are set parameter of item $i^{th}$, $\widehat{p_{ij}}$ as the average probability of the correct answer item $i^{th}$ of the examinee's group $j^{th}$ (have the average ability $\widehat{\theta_j}$) and is a result of the average ability of the examinee's group $j^{th}$ for item parameters in a bank and can be calculated using statistical method of CTT [8] as the number of correct answers divided by the total number of assigned items of the examinee's group $j^{th}$, belonging to the interval of examinee's ability (*min, Max*):

$$\widehat{\theta_J} = min + (|min| + |Max|)\sum_{t=1}^{K}\frac{L_t}{N} \tag{13}$$

where *min* and *max* are minimum and maximum value of examinee's ability. In this paper, authors give *min* = -3 and *max* = 3 because examinee's ability depends on the value of information function Eq. (2). Baker [16] mentioned that, if the amount of information is large, it means that an examinee whose true ability is at that level can be estimated with precision; i.e., all the estimates will be reasonably close to the true value and show that with range (-3, 3), examinee's ability is estimated with some precision. Outside this range, the amount of information decreases rapidly, and the corresponding examinee's ability is not estimated very well. $K$ is the total number of the examinee's group $j^{th}$ who has the average ability $\widehat{\theta_J}$, $L_t$ is the number of correct answers of examinee $t^{th}$ in the examinee's group $j^{th}$, and $N$ is the total number item giving an estimation for the examinee's group $j^{th}$.

To find the smallest value of the function in the Eq. (12) to calculate the parameter value of each item $i^{th}$. In the general case, we need to solve the following system of equations:

$$\begin{cases} \frac{dg(a_i,b_i,c_i)}{da_i} = 0 \\ \frac{dg(a_i,b_i,c_i)}{db_i} = 0 \\ \frac{dg(a_i,b_i,c_i)}{dc_i} = 0 \end{cases} \tag{14}$$

Likewise, the problem of finding the values of the items parameter under the proposed model. In this instance, not one but two functions are minimized:

$$\begin{aligned} g_0(a0_i, b0_i, c_i) &= \sum_{j=1}^{M}\begin{pmatrix} p\left(u^{(k)} = 1|u^{(k-1)} = 0, \widehat{\theta_J}, a0_i, b0_i, c_i\right) \\ -\widehat{p_{\iota J}}\left(u^{(k)} = 1|u^{(k-1)} = 0\right) \end{pmatrix}^2 \to min \\ g_1(a1_i, b1_i, c_i) &= \sum_{j=1}^{M}\begin{pmatrix} p\left(u^{(k)} = 1|u^{(k-1)} = 1, \widehat{\theta_J}, a1_i, b1_i, c_i\right) \\ -\widehat{p_{\iota J}}\left(u^{(k)} = 1|u^{(k-1)} = 1\right) \end{pmatrix}^2 \to min \end{aligned} \tag{15}$$

Hooke and Jeeves [17] offers the direct search method, due to the complexity of Eq. (14), the smallest value of the function in the Eqs. (12) and (15). One of the critical advantages of this method allows search from an arbitrary basis vector. The finding is a sequence of search steps by coordinates around the basis vector to reach the vector that has the objective function value is smaller. If successful, it will transfer the basic vector to the better one found. We call it optimal vector and keep going on that direction to the sample vector, then search by coordinates around the sample vector. If the search was the better vector, then continued searching around the new template vector. If not successful, it will go back to the previous basis vector or reduce the length of the step search.

This method was developed to solve the problems of optimization (in this case is looking for the smallest value of a function) without concern about the limit. However, to reduce the search time, the efficiency of search engines should be improved in the reassessment of item parameters. We need to supplement the limited value range of each item parameter. In this study, the authors propose the limited value range of item difficulty *b* and discrimination a as follows: $-3 < b < 3$ and $0 < a < 2$ [11].

The direct search method of Hooke-Jeeves innovated by putting more on the limited value range for item parameters. It will help in finding the minimum value of the objective function is more advantageous. According to the proposed method, when a new value of item parameters is found, then you have to consider whether that value is located in the limited value range. If it is true, the objective function is calculated, as usual. Otherwise, the objective function will get great value. Therefore, finding the optimum value of item parameters will be made in the region of limited value. The assessment algorithm of item parameters described in Fig. 4.

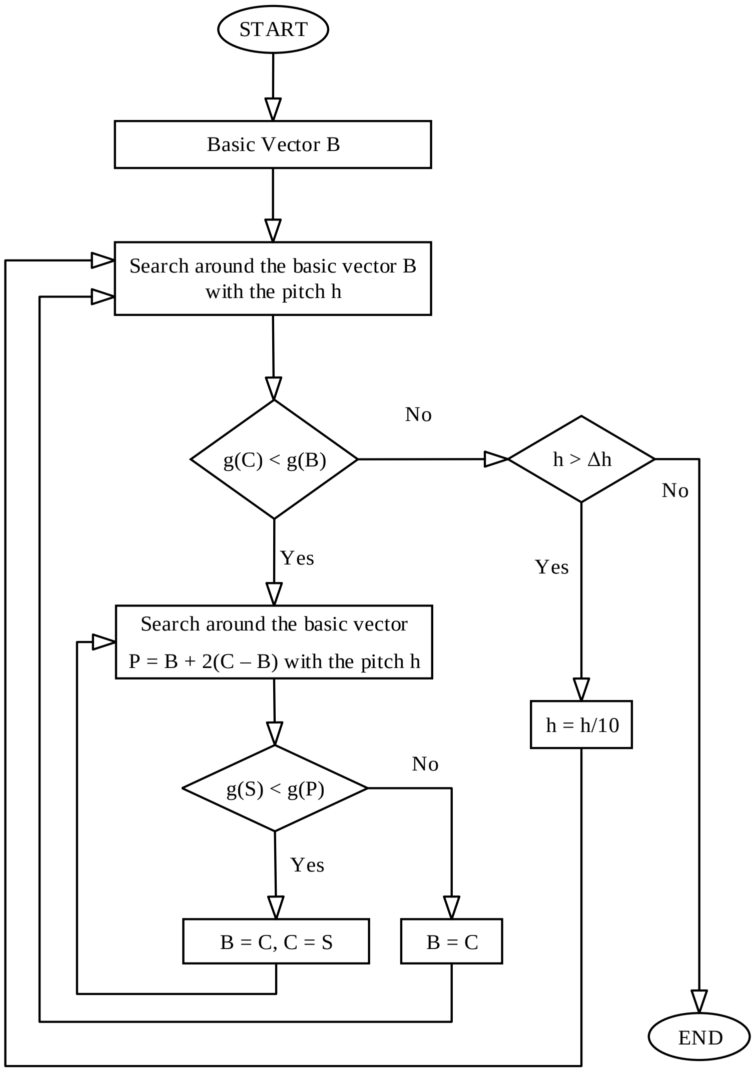


**Fig. 4. Parameter evaluation algorithm.**

The proposed algorithm is started by selecting an initial basis vector $B$ (vector of item parameters) and the pitch $h$ for each of the parameters need to be estimated:

- **Step 1**: Search for an optimal vector around the basis vector $B$ in the change direction of each parameter, calculate the value of objective function with the set of vectors: $\{C_1(B + he1), C_2(B - he1), C_3(B + he2), C_4(B - he2)\}$, in which, $e1$, $e2$ respectively that the unit vector in the direction of the item difficulty $b$ and the item discrimination $a$: if $g(C_i) < g(B)$ then assign the optimal vector $C = C_i$ and move on step 3 with $i = 1, \ldots, 4$. In the opposite case, $C = B$ and move on step 2.
- **Step 2**: If $C = B$ then continue to reduce the value of the objective function as step 1 around the basis vector $B$ but with pitch value h decreased. In this article, the author proposes the pitch value $h$ will decrease ten times corresponding to each reducing time. Continue to perform it until the pitch value h reached the small value $\Delta h$. In this article, the authors offer $\Delta h = 0.01$.
- **Step 3**: If $C \neq B$ then searching around the new basis vectors $P = B + 2(C - B)$ in the direction $(C - B)$. Vector $P$ achieved by moving from $B$ to $C$ and continue in an equal distance in the same direction.
- **Step 4**: Continue the search for optimal vector around the basis vector $P$ in the direction based on the change of the item parameters with the pitch value $h$. At that time, we will find the optimal vector $S$, If $S = P$, then switch to step 5. Otherwise, go back to step 6.
- **Step 5**: assign the basis vector value $B = C$ and go back to step 1.
- **Step 6**: assigns the vector value $B = C$, $C = S$ and go back to step 3.

This method is modified to account for restrictions. Clearly, during the implementation, it is checked whether each point obtained in the search process belongs to the constraint area. If true, then the target solution is calculated in the usual way. If false, the target function is assigned an immense value. Thus, the search is performed again in the permissible region in the direction of the minimum point inside this region.

## 3. Experiment

The authors start to build the website with the use-case diagram as shown in Fig. 5. The system architecture consists of three main modules: Interface Agent module, Adaptive Testing module, and Estimation Item module. In which, the Interface Agent module enables the system to interact with the users, the Adaptive Testing module allows for the assessment of the examinee's ability with two algorithms based on the proposed model and the Birnbaum model, and the Estimation Item module is used to evaluate item parameters. The authors have carried out building the item bank of Object-Oriented Programming (OOP) subject - faculty of Information Technology, University of Danang, University of Science and Technology. Item of the item bank uses the Multiple Choice Question with the number of answers is 4 in each. Therefore, the guessing parameter will have the same value, equal to 0.25, in both the Birnbaum models and the proposed model. The processing of the received statistical information on the responses was carried out with the help of module Estimation item and analysis of the statistical information described above. As a result of data processing, an assessment of the item parameters was given, which is shown in Table 1.

**Table 1. Result of item parameters of OOP subject with both Birnbaum model and the proposed model.**

| Item | Birnbaum model | | Proposed model | | | |
|---|---|---|---|---|---|---|
| | ***a*** | ***b*** | $a_0$ | $a_1$ | $b_0$ | $b_1$ |
| **1** | 0.517 | -0.936 | 0.944 | 1.545 | 0.538 | -2.978 |
| **2** | 1.837 | -2.911 | 0.498 | 1.899 | 0.123 | -2.954 |
| **3** | 0.519 | -2.617 | 0.478 | 1.911 | 0.159 | -2.858 |
| **4** | 1.678 | 0.166 | 1.932 | 1.268 | 0.846 | -2.675 |
| **5** | 1.549 | -0.541 | 0.447 | 1.638 | 0.815 | -2.655 |
| **6** | 0.552 | -1.534 | 0.607 | 1.127 | 0.815 | -2.296 |
| **7** | 0.494 | -2.121 | 0.527 | 0.549 | -2.149 | -2.127 |
| **8** | 0.657 | -2.154 | 0.421 | 1.822 | -2.036 | -2.031 |
| **9** | 0.714 | -2.835 | 1.955 | 0.556 | -2.931 | -1.347 |
| **10** | 0.749 | -1.414 | 1.872 | 0.500 | -1.265 | -0.979 |
| **11** | 0.834 | -0.936 | 0.726 | 0.802 | -0.912 | -0.857 |
| **12** | 0.788 | -1.629 | 2.200 | 0.500 | -2.415 | -0.634 |
| **13** | 0.441 | -0.612 | 0.673 | 0.502 | -0.616 | -0.588 |
| **14** | 1.734 | -0.098 | 1.641 | 1.756 | -0.098 | -0.127 |
| **15** | 1.842 | -0.216 | 1.989 | 1.693 | -0.128 | -0.077 |
| **16** | 0.462 | -1.733 | 1.919 | 0.715 | -2.603 | -0.040 |
| **17** | 0.423 | -2.258 | 1.832 | 0.505 | -2.888 | -0.004 |
| **18** | 0.406 | -1.112 | 1.858 | 1.869 | -2.296 | 0.218 |
| **19** | 1.452 | 0.443 | 0.418 | 1.728 | 0.007 | 0.321 |
| **20** | 1.771 | 0.307 | 1.549 | 1.823 | 2.791 | 0.433 |
| **21** | 0.580 | 0.124 | 0.517 | 1.949 | -0.761 | 0.487 |
| **22** | 1.521 | 2.934 | 1.803 | 1.147 | 2.888 | 0.542 |
| **23** | 1.647 | 1.517 | 1.736 | 1.007 | 1.578 | 0.565 |
| **24** | 1.162 | 0.872 | 0.579 | 1.634 | 2.446 | 0.588 |
| **25** | 0.370 | -1.774 | 1.723 | 0.680 | -2.594 | 0.600 |
| **26** | 1.957 | 0.839 | 1.823 | 1.936 | 0.716 | 0.764 |
| **27** | 0.528 | 0.953 | 1.897 | 1.732 | -2.407 | 0.841 |
| **28** | 1.784 | 1.023 | 1.851 | 1.624 | 2.971 | 0.847 |
| **29** | 1.594 | 1.278 | 1.657 | 1.032 | 2.695 | 0.848 |
| **30** | 0.593 | -2.218 | 0.467 | 1.222 | -2.015 | 1.401 |
| **31** | 0.753 | 2.856 | 1.548 | 1.892 | -0.066 | 1.451 |
| **32** | 0.584 | 2.704 | 1.783 | 1.605 | 0.541 | 1.489 |
| **33** | 1.827 | 2.401 | 0.423 | 1.886 | 0.080 | 1.503 |
| **34** | 1.354 | 1.495 | 1.426 | 1.739 | -2.678 | 1.541 |
| **35** | 1.783 | 2.254 | 1.953 | 1.542 | -2.633 | 1.543 |
| **36** | 1.523 | 1.108 | 1.892 | 1.288 | -0.865 | 1.568 |
| **37** | 0.578 | 1.380 | 1.864 | 1.785 | -1.947 | 1.588 |
| **38** | 1.271 | 2.041 | 1.405 | 1.618 | 1.578 | 1.675 |
| **39** | 1.696 | 1.897 | 0.511 | 1.921 | 1.279 | 1.768 |
| **40** | 0.616 | -0.768 | 0.597 | 1.863 | -2.151 | 2.872 |

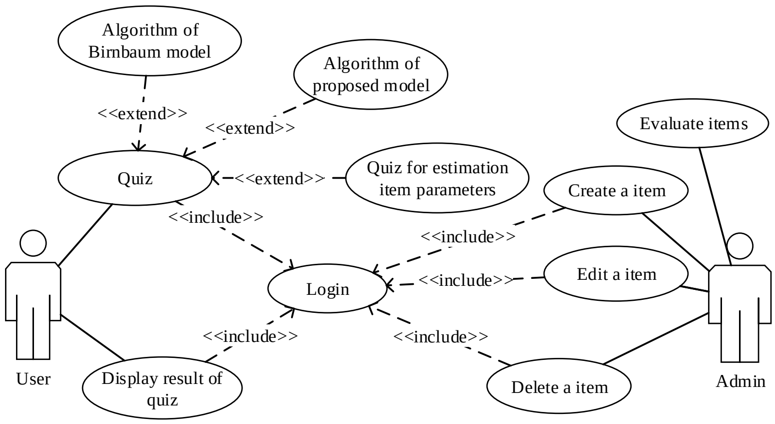


**Fig. 5. Use-case diagram of website.**

## 4. Results and Discussion

The authors assessed the examinee's ability of the two groups regarding the OOP subject based on the proposed model and the Birnbaum model. Examinees in the two groups had the same ability level responding well to the OOP subject. The consequence of the quiz process is shown in Table 2 and Fig. 6.

According to Fig. 6, the examinee's ability in both models nearly gave the same outcome; the examinees in the two groups presented a significant level of capacity on the OOP subject based on the given item bank. Thus, the proposed model answers the criteria for accurately assessing the examinee's ability. Besides, the Fig. 6 shows that the average number of assigned items used to evaluate the ability of examinees who following the proposed model is 19.5 items, while the same figure under the Birnbaum model is 23 items. Furthermore, Fig. 6 also describes that the various ranges of the examinee's ability according to the proposed model are smaller than the one in the Birnbaum model. Hence, it is concluded that adaptive testing on the base of the proposed model is more efficient than the Birnbaum model of IRT.

The correlation analysis of the sequence of answers carried out within the framework of this work allowed us to propose the modernization of the Birnbaum model by the mathematical apparatus of Markov chains. The application of this mathematical apparatus will enable us to obtain a better adequate testing model, which takes into account the dependence of the response of the tested person on the result obtained in the previous step. Such a model, in the author's opinion, more adequately describes the testing process when displaying the results of answers to questions at each step of testing. In this article, to assess the examinee's ability, it is suggested to use the method of maximum likelihood known in mathematical statistics. The variant of using this method is given for both the Birnbaum model and for the proposed model.

**Table 2. Assessment result of the examinee's ability on OOP subject for both models (Birnbaum model and proposed model).**

| Model CAT | | Number of assigned items | Average number of assigned items |
|---|---|---|---|
| **Proposed model** | Examinee 1 | 18 | 19.5 |
| | Examinee 2 | 21 | |
| **Birnbaum model** | Examinee 3 | 24 | 23 |
| | Examinee 4 | 22 | |

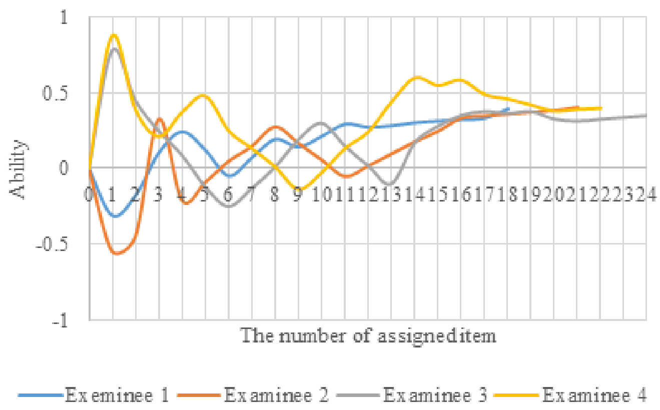


**Fig. 6. Adaption of examinee.**

## 5. Conclusions

In this work, the Birnbaum model was selected for developing CAT. Nevertheless, the Birnbaum model of IRT does not take into account the relationship between the sequences of responses. Thus, the authors proposed model adaptive testing based on the Birnbaum model used the mathematical apparatus of Markov chains to enhance the effectiveness and the accuracy of the adaptive testing process.

Besides, the set of item parameters also influence the accuracy of adaptive testing models. To solve the problem of estimating the item parameters, the direct search Hooke-Jeeves method with a limit range of item parameters is deployed in both the Birnbaum model and the proposed model. The method of evaluating item parameters allows selecting the proper set of items for the assessment of the examinee's ability during the adaptive testing process.

### Nomenclatures

| | |
|---|---|
| $a_i$ | An item discrimination parameter of item $i^{th}$ |
| $b_i$ | An item difficulty parameter of item $i^{th}$ |
| $c_i$ | An item pseudo-guessing parameter of item $i^{th}$ |
| $p_{ji}$ | A probability of a correct answer to item $i^{th}$ of $j^{th}$ examinee |
| $I_i$ | Information function value of item $i^{th}$ |

***Greek Symbols***

| | |
|---|---|
| $\theta_j$ | Ability of $j^{th}$ examinee |

**Abbreviations**

| | |
|---|---|
| CAT | Computerized Adaptive Testing |
| CTT | Classical Test Theory |
| IRT | Item Response Theory |
| OOP | Object Oriented Programming |

## References

1. Novick, M.R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, 3(1), 1-18.
2. van-der-Linden, W.J.; and Glas, C.A.W. (2010). *Elements of adaptive testing*. New York, United States of America: Springer Science+Business Media.
3. Meijer, R.R.; and Nering, M.L. (1999). Computerized adaptive testing: Overview and introduction. *Applied Psychological Measurement*, 23(3), 187-194.
4. Chen, C.-M.; Lee, H.-M.; and Chen, Y.-H. (2005). Personalized e-learning system using item response theory. *Computers & Education*, 44(3), 237-255.
5. Wauters, K.; Desmet, P.; and Noortgate, W.V.d. (2010). Adaptive item-based learning environments based on the item response theory: Possibilities and challenges. *Journal of Computer Assisted Learning*, 26(6), 549-562.
6. van-der-Linden, W.J.; and Glas, C.A.W. (2000). *Computerized adaptive testing: Theory and practice*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
7. Crocker, L.M.; and Algina, J. (1986). *Introduction to classical and modern test theory*. New York, United States of America: Holt, Rinehart and Winston Inc.

8. Bichi, A.A.; Embong, R.; Mamat, M.; and Maiwada, D.A. (2015). Comparison of classical test theory and item response theory: A review of empirical studies. *Australian Journal of Basic and Applied Sciences*, 9(7), 549-556.
9. Lord, F.M. (1980). *Applications of item response theory to practical testing problems*. Mahwah, New Jersey, United States of America: Lawrence Erlbaum Associates.
10. Weiss, D.J. (2014). *New horizon testing: Latent trait test theory and computerized adaptive testings*. New York, United States of America:Academic Press, Inc.
11. Baker, F.B.; and Kim, S.-H. (2004). *Item response theory: Parameter estimation techniques* (2nd ed.). New York United States of America: Marcel Dekker Inc.
12. Pathak, A.; Patro, K.; Pathak, M.; and Valecha, M. (2013). Item response theory. *International Journal of Computer Science and Mobile Computing (IJCSMC)*, 2(11), 7-11.
13. Hambleton, R.K.; and Swaminathan, H. (2013). *Item response theory: Principles and applications*. Massachusetts, United States of America: Kluwer Academic Publishers.
14. Sericola, B. (2013). *Markov chains: Theory and application*. Hoboken, New Jersey, United States of America: John Wiley & Sons, Inc.
15. Samuel, K.; Balakrishnan, N.; Read, C.B.; Vidakovic, B.; and Norman, L.J. (2006). *Encyclopedia of statistical sciences* (2nd ed.). Hoboken, New Jersey, United States of America: Wiley-Interscience.
16. Baker, F. (2001). *The basic of item response theory* (2nd ed.). Washington, D.C., United States of America: ERIC Clearinghouse on Assessment and Evaluation.
17. Hooke, R.; and Jeeves, T. (1961). Direct search solution of numerical and statistical problems. *Journal of the ACM*, 8(2), 212-229.