Application of Markov Chain for Comparing the Performance of Students in Adaptive E-Assessment

V.Geetha, Professor, Department of Computer Applications, Kumaraguru College of Technology, Coimbatore

G.S.Nandakumar, Associate Professor, Department of Computer Science & Engineering, Kumaraguru College of Technology, Coimbatore

R.Vivek, Department of Computer Applications, Kumaraguru College of Technology, Coimbatore

Abstract - Information and Communication Technologies have become an integral part of teaching learning process. New Technologies have provided the educational field with innovations to improve the teaching learning process. One of the fastest evolving fields in the teaching learning process is assessing the performance of the students. Adaptive E-assessment with multiple choice questions is one of the reliable methods of assessing the students on a large scale. Its success largely depends on the caliber of the questions and their classification to different degree of difficulty to suit individual student ability. An adaptive E-assessment approach has been articulated and this study applies Markov chain to assess the consistency of grading the questions and its effectiveness in the assessment of student performance. Comparisons of performance between two batches of students with different backgrounds have also been carried out.

Keywords: Adaptive E-assessment; Markov model; Multiple Choice Questions

I. INTRODUCTION

Assessment is an integral part of education associated with learning outcomes. The goal of the assessment is to estimate the knowledge that has been acquired by the students through learning. Technology plays a vibrant role in supporting the assessment of knowledge in the learning process. The development of new digital technologies has facilitated the implementation of web-based assessment tools. Adaptive E-assessment is a popular mode of knowledge assessment used by the educational institutions. Adaptability is the significant functionality in this assessment, in which the questions are selected judiciously to fit the students’ level of knowledge, resulting in generating different sets of questions for each student. The difficulty level of the questions adjusts dynamically depending on the individual’s answering pattern. This has great prospective to make the learning environment more realistic to the learners. The mastery of a student will be assessed with lesser questions than in conventional assessments [1][2]. These systems can determine a student’s score with less number of questions which reduced the length of the test by atleast by 60%, resulting in an improvement in examinees’ performance [3].

Multiple choice questions are commonly used in adaptive assessments. A huge question bank is needed to target the different categories of students with different skills. Automated assessment systems generate a great deal of information and there is a need of leveraging this data into knowledge which can then be used to analyze the performance of the learners.
A conceptual design for an adaptive assessment tool has been developed using client server platform [4]. Reference [5] in her doctoral work has explored the problems associated with the use of computer aided assessment in higher education of U.K. An E-Assessment model has been implemented and discussed in [6]. Markov chain is a stochastic modeling technique that permits transitions from one state to another using some probabilistic rules. The important characteristics of Markov chain is that the probability of arriving at the present state depends only on the previous state and not on any of the preceding states. Markov chains are used in many real world applications including education. Some of the applications which use Markov chain have been discussed in [7]. Markov Chain has been used for predicting the behavior of customers by a telecom service provider[8]. Markov chain has been applied in educational planning to indicate the average time taken by students to stay in a tertiary institution [9]. The website browsing patterns have been modeled using Markov chains and an analysis has been carried out [10]. The navigation behavior of students in a web based E-learning system of an educational institution have been exhibited using Markov chain to determine the critical periods of site navigation [11]. Markov analysis has been carried out to find out the possibility of students dropping out from a university has been analyzed using Markov chains [12].

The student movement in a developed academic institution has been explored using Markov analysis [13]. Hidden Markov models have been used to prototype the activities of school students while using an intelligent tutoring system [14]. The outcome of student learning in an e-has been learning environment using Hidden Markov models has been studied [15]. Homogenous Markov chains were used to define the effect of teaching and learning procedures in educational institutions [16]. Markov chains have been used to evaluate the authenticity of the questions’ classification. Students have been clustered based on their performance in answering the questions during a specified time duration [17].

ILADATIVE E-ASSESSMENT STRATEGY

An adaptive E-assessment approach has been developed using PHP and MySQL to test the ability of students in ‘C’ programming language. Multiple-Choice Questions (MCQ) are extensively used in adaptive assessments.

A. Question Bank Creation

The quality of the questions stored in the question bank determines the success of the adaptive test. The total number of questions in the bank must be sufficient to supply the needed questions throughout the assessment period. This helps to achieve high accuracy rate during the measurement range. This criterion predominantly states that there should be sufficient number of calibrated questions at all the difficulty levels. The question bank is calibrated to identify the measurement traits of large E-Assessment. It basically assigns a difficulty level to each question stored in the question bank. A good question bank will consist of adequate number of suitable and well organized questions to test the different levels of proficiency. The difficulty of the questions should be extensive enough to cover the wide scope of the students’ calibre. The chances of each student getting different set of questions are greatly enriched in having a large question bank.

Each question in the question bank is assigned a difficulty level known as Degree of Toughness (DT). The questions are categorized into five DT levels from 1(Basic) to 5(Very difficult). The interesting feature of this strategy is that the student can initially choose the DT of the questions as soon as he starts the system of examination. If he opts for the n\textsuperscript{th} DT (n=1,2,3, 4, 5), questions will be randomly displayed from the question bank for the chosen DT, for which the student answers.

The algorithm covering the four different cases of inputs is discussed below:

Case 1: If the candidate answers the first three questions of the n\textsuperscript{th} DT correctly, the system will shift to the next higher (n+1) DT provided n ≠ 5. If n= 5, the system continues to display the questions from the same level.

Case 2: In case the candidate answers the first three questions of the n\textsuperscript{th} DT incorrectly, the system will shift to next lower (n-1) DT provided n ≠ 1. If n = 1, the system continues to display from the 1\textsuperscript{st} DT regardless of the number of wrong answers given.

Case 3: This case discusses a situation where the candidate answers either one or two questions correctly out of the first three questions from the n\textsuperscript{th} DT. The system displays one more question from the n\textsuperscript{th} DT. The examinee encounters a total of four questions at the n\textsuperscript{th} DT. A total of three correct answers moves to the next higher (n+1) DT, provided n ≠ 5 and a total of three wrong answers moves to the next lower (n-1) DT, provided n ≠ 1.

Case 4: If the candidate answers two questions correctly out of the first four questions from the n\textsuperscript{th} DT, one more question from the same DT is displayed. A total of three correct answers out of five for the
given questions, shifts to the next higher \((n+1)\) DT; otherwise to next lower \((n-1)\) DT. Nevertheless, moving to a lower or higher DT does not take place when \(n=1\) or \(n=5\) respectively.

The difficulty level of a question has to be revised at periodic intervals, after a wide range of students undergo the tests and the question has been asked adequately large number of times.

**B. Evaluation Procedure**

The grades for a question in each DT are given in Table I. It can be seen from the table that marks accelerate corresponding to the DT.

<table>
<thead>
<tr>
<th>DT Level</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mark</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
</tr>
</tbody>
</table>

The total marks and the duration of the examination can be modified to suit the needs of the individual subject. The test will get completed either at the expiration of the time frame or the examinee has undertaken questions for the assigned maximum marks, whichever occurs first. The marks scored in the test and the number of DT – wise questions asked and answered will be displayed at the end of the test.

**III. MARKOV MODEL FOR ADAPTIVE E-ASSESSMENT**

Markov models were applied to classify and compare the performance of two batches of students with different background. The students have been classified into five groups based on the steady state reached in the assessment. The DT levels and their associated grades are shown in Table II.

<table>
<thead>
<tr>
<th>DT level</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT1</td>
<td>Far below average</td>
</tr>
<tr>
<td>DT2</td>
<td>Below average</td>
</tr>
<tr>
<td>DT3</td>
<td>Average</td>
</tr>
<tr>
<td>DT4</td>
<td>Bright</td>
</tr>
<tr>
<td>DT5</td>
<td>Very Bright</td>
</tr>
</tbody>
</table>

A candidate can navigate across various DT levels in a gradual manner which can be characterized as a series of random variables for each student, representing the state of the system DT₁, DT₂, ... DT₅. A candidate beginning the test at DT level \(k\), with \(k\) ranging from 1 to 5, can either move to DT₁, DT₂, or DT₃, or at each step. It is not possible to move to any other levels. However, moving to a higher or lower DT does not take place when \(k=5\) or \(k=1\) respectively. The probability of moving from one state to the same state is denoted by \(P_{ii}\) and will be equal to zero for \(k = 2, 3 \& 4\). This follows Markov model because each state depends only on the previous state. The transition between the states can be seen as a Markov chain and the transition probability matrix is depicted in (1).

\[
P = \begin{bmatrix}
P_{11} & P_{12} & 0 & 0 & 0 \\
P_{21} & 0 & P_{23} & 0 & 0 \\
0 & P_{32} & 0 & P_{34} & 0 \\
0 & 0 & P_{43} & 0 & P_{45} \\
0 & 0 & 0 & P_{54} & P_{55}
\end{bmatrix}
\]

(1)

The initial state transition probability for adaptive assessment of 1st batch students is shown in (2).

\[
P = \begin{bmatrix}
0.006 & 0.994 & 0 & 0 & 0 \\
0.368 & 0 & 0.632 & 0 & 0 \\
0 & 0.5 & 0 & 0.5 & 0 \\
0 & 0 & 0.56 & 0 & 0.44 \\
0 & 0 & 0 & 0.875 & 0.125
\end{bmatrix}
\]

(2)

When the assessment was conducted for the 2nd batch students, the transition probabilities resulted as shown in (3).

The \(n\)-step probabilities are calculated to predict the future states based on the current state. The \(n\)-step transition probability of a Markov chain is the probability of transitioning from state \(i\) to state \(j\) in \(m\) transitions. It gives the conditional probability of the candidates starting at DT₁, will reach the other DT levels in \(m\) transitions.

\[
P = \begin{bmatrix}
0.043 & 0.957 & 0 & 0 & 0 \\
0.636 & 0 & 0.364 & 0 & 0 \\
0 & 0.488 & 0.049 & 0.463 & 0 \\
0 & 0 & 0.625 & 0 & 0.375 \\
0 & 0 & 0 & 0.935 & 0.065
\end{bmatrix}
\]

(3)
Table III shows the percentage of students who start their test at each DT level. It can be observed that more number of students start at DT1.

<table>
<thead>
<tr>
<th>Difficulty level</th>
<th>Batch 1</th>
<th>Batch 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT1</td>
<td>78.35</td>
<td>30.5</td>
</tr>
<tr>
<td>DT2</td>
<td>8.23</td>
<td>7.3</td>
</tr>
<tr>
<td>DT3</td>
<td>6.06</td>
<td>25.8</td>
</tr>
<tr>
<td>DT4</td>
<td>3.9</td>
<td>15.9</td>
</tr>
<tr>
<td>DT5</td>
<td>3.46</td>
<td>20.5</td>
</tr>
</tbody>
</table>

The transition probability matrix after 10 transitions was calculated to observe the DT level of candidates in the adaptive assessment for both the batches and is shown in Tables IV & V respectively. In both the situations, it can be seen that probabilities of all DT levels remain greater than zero after 5 transitions, which clearly shows that the Markov chain is ergodic i.e. all states are reachable from every other state over a period of time. A student starting in a DT level will be able to reach other DT levels during the course of the test. This clearly indicates that the questions have been classified correctly in the appropriate DT levels.

In Table IV, it can be seen that out of the students who have started in DT1, 15.18% stay in DT1, 5.08% moved to DT2, 50.66% moved to DT3, 8.87% moved to DT4 and 20.1% moved to DT5. Of the students who started in DT2, 0.99% could reach DT5. Among the students who started in DT5, 16.03% moved down to DT3.

<table>
<thead>
<tr>
<th>Percentage of students starting at DT1 moving to other DT levels</th>
<th>Batch 1</th>
<th>Batch 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT1</td>
<td>0.1518</td>
<td>0.0508</td>
</tr>
<tr>
<td>DT2</td>
<td>0.0160</td>
<td>0.3965</td>
</tr>
<tr>
<td>DT3</td>
<td>0.1358</td>
<td>0.0606</td>
</tr>
<tr>
<td>DT4</td>
<td>0.0338</td>
<td>0.2917</td>
</tr>
<tr>
<td>DT5</td>
<td>0.0739</td>
<td>0.1603</td>
</tr>
</tbody>
</table>

In Table V, it can be observed that out of the students who have started in DT1, 15.18% stay in DT1, 5.08% moved to DT2, 50.66% moved to DT3, 8.87% moved to DT4 and 20.1% moved to DT5. Of the students who started in DT2, 0.99% could reach DT5. Among the students who started in DT5, 16.03% moved down to DT3.

In batch 2, out of the students who start in DT1, 39.17% move to DT3 and 36.89% stay at DT1. Of the students started in DT2, 57.47% of the students remain in DT2 while 30.6% move to DT3. Out of students who start in DT3, 41.47% stay in DT3, while 36.65% moved down to DT1. Fig 1 shows the comparison between the percentages of students starting at DT1 in both the batches.

In most of the cases, it can be observed that the percentage of students moving to higher DT levels is more in batch 1 than in batch 2. Batch 2 students starting at higher DT levels tend to move more towards lower DT levels in course of time.

Thus Markov models are used to compare the performance of students using the transition between the DT levels rather than using the score. Guessing the answers in adaptive test might affect the final score, but not on the DT level over a period of time.

A. Steady State Probability Matrix

The probability of the system being in particular state after a large number of stages is called steady state probability.
Steady state probabilities can be found by solving the set of equations (5) and (6).

\[ \mu_j = \sum_{i=0}^{K} \mu_i P_{ij} \quad \text{for } j = 0, 1, 2, \ldots, K \]  

(5)

\[ \sum_{j=0}^{K} \mu_j = 1 \]  

(6)

where \( \mu = [\mu_1, \mu_2, \ldots, \mu_5] \) is a unique probability vector and \( K \) is the number of states.

The probability of starting at state \( j \) (\( j = 1, 2, 3, 4, 5 \)) in the long run will settle at values that are solutions of equation \( \mu P = \mu \) and represented in matrix notation as given in (7).

\[
\begin{bmatrix}
0.06 & 0.94 & 0 & 0 & 0 \\
0.38 & 0 & 0.63 & 0 & 0 \\
0 & 0.5 & 0 & 0.5 & 0 \\
0 & 0 & 0.56 & 0 & 0.44 \\
0 & 0 & 0 & 0.87 & 0.125
\end{bmatrix}
\]

(7)

By expanding the above matrix, the linear sets of equations from (8) to (13) are obtained.

\[ 0.006\mu_1 + 0.368\mu_2 = \mu_1 \]  

(8)

\[ 0.994\mu_1 + 0.5\mu_3 = \mu_2 \]  

(9)

\[ 0.632\mu_2 + 0.5\mu_4 = \mu_3 \]  

(10)

\[ 0.5\mu_3 + 0.875\mu_5 = \mu_4 \]  

(11)

\[ 0.44\mu_4 + 0.125\mu_5 = \mu_5 \]  

(12)

and

\[ \sum_{i=1}^{5} \mu_i = 1 \]  

(13)

The following values have been obtained by solving (8)-(13), \( \mu_1=0.087, \mu_2=0.234, \mu_3=0.296, \mu_4=0.234 \) and \( \mu_5=0.148 \). The probability of students remaining in DT1 is 0.087, DT2 is 0.234, DT3 is 0.296, DT4 is 0.234 and DT5 is 0.148. Steady state probabilities indicate that a large section of students reach DT3 followed by DT5 and DT4.

Similarly for batch 2, matrix notation for steady state probability is \( \mu P = \mu \) and represented in matrix notation as given in (14).

\[
\begin{bmatrix}
0.043 & 0.957 & 0 & 0 & 0 \\
0.636 & 0 & 0.364 & 0 & 0 \\
0.513 & 0 & 0.487 & 0 & 0 \\
0 & 0 & 0.625 & 0 & 0.375 \\
0 & 0 & 0 & 0.935 & 0.065
\end{bmatrix} = [\mu_1, \mu_2, \mu_3, \mu_4, \mu_5] \]  

(14)

and we have the linear set of equations from (15) to (20).

\[ 0.043\mu_1 + 0.636\mu_2 = \mu_1 \]  

(15)

\[ 7\mu_1 + 0.513\mu_3 = \mu_2 \]  

(16)

\[ 0.364\mu_2 + 0.625\mu_4 = \mu_3 \]  

(17)

\[ 0.487\mu_3 + 0.935\mu_5 = \mu_4 \]  

(18)

\[ 0.37 \mu_4 + 0.065\mu_5 = \mu_5 \]  

(19)

\[ \sum_{i=1}^{5} \mu_i = 1 \]  

(20)

By solving the above equations, we obtain \( \mu_1=0.211, \mu_2=0.317, \mu_3=0.225, \mu_4=0.176 \) and \( \mu_5=0.071 \). This indicates that the probability of students remaining in DT1 is 0.211. DT2 is 0.317, DT3 is 0.226, DT4 is 0.176 and DT5 is 0.071. The steady state probabilities of both the batches are graphically represented in Fig. 2.

![Fig. 2. Steady state probabilities for both the batches](image-url)

Steady state probabilities indicate that a large section of students reach DT2 followed by DT3 and DT1. The relative performance between both the batches is compared in Fig. 2. It can be seen that more number of students reach higher DT levels in Batch 1 than in Batch 2. A large section of students stay at DT3 in batch 1 whereas it is in DT2 in Batch 2.
V. CONCLUSION

Markov chains have been used to demonstrate the authenticity of grading the questions in adaptive assessment. The work also concentrated on using Markov modeling to classify the performance of students based on the DT levels attained instead of using the score obtained in the assessment. This is a better approach for analyzing the performance of students because guessing the answer for the questions will impact the score but not the DT levels. A comparison between two groups of students with different backgrounds has been made and results were analyzed.

REFERENCES
