Automated Exam Question Generator using Genetic Algorithm

Tengku Nurulhuda Tengku Abd Rahim
Cradmetech Sdn. Bhd,
Temasya Industrial Park, 40150 Shah Alam, Selangor, Malaysia

Zalilah Abd Aziz, Rose Hafsah Ab Rauf, and Noratikah Shamsudin
Faculty of Computer and Mathematical Sciences,
Universiti Teknologi MARA,
40450 Shah Alam, Selangor, Malaysia
zalilah@tmsk.uitm.edu.my

Abstract—Manual preparation of exam questions is a challenging task for educators especially within a short time frame. It requires a lot of time and efforts in order to meet the standard quality of exam questions. This research introduces an automated exam question generator to resolve this issue in preparation of multiple choice exam questions. The generator can auto generate new exam questions set using Genetic Algorithm and covers six levels of Bloom's Taxonomy to produce high quality exam questions that can evaluate different level of learners based on Bloom's cognitive domains and the selection of chapters made by educators. The prototype with 500 sample questions has been run 50 times with different number of chapters selected for each test case. It manages to achieve 90% for the highest exam questions weightage while the average value of exam questions weightage percentage generated is 70%. The lowest exam questions weightage percentage generated is 40%. The result is affected by the smaller number of questions for each Bloom's taxonomy level in questions bank. The automated exam generator can extends to be used for any type of exam questions and it can be used for preparation of quiz or test questions.

Keywords—Automated Exam Questions Generator; Bloom's Taxonomy; Artificial Intelligence; Genetic Algorithm; Cognitive; Random Search.

I. INTRODUCTION

A. Overview

Exam is a part of evaluation given by educators to their students in order to assess their students' performance based on their learning outcome. A good quality of exam questions can determine good quality of students and may also reflect the quality work of educators and their educational institutions. A good exam paper should consist of various difficulty levels to tolerate the different capability of students. Manual preparation of exam questions requires a lot of work to be done to make sure all the guidelines given by educational institutions are followed by educators during preparation of exam questions. It requires lots of educators' efforts and time since it is tedious and meticulous which can sometimes leads to human mistake. Common issues that educators face in preparing exam questions are time constraint to prepare good exam questions based on institution's examination guideline, predictive question pattern where the same question style repetitively produced over the time [1], bias exam questions either it is too easy or too difficult for the students [1] and exam question which does not cover all six levels of Bloom's Taxonomy's cognitive domain. Thus, an automated exam question is a must to reduce burden on educators in preparing the exam questions manually.

Sponsored by Universiti Teknologi MARA (UiTM), Malaysia

The purpose of this research is to ease the educators work in the process of preparation exam questions and giving them more time to concentrate on teaching materials and strengthen their teaching techniques without being burdened with the exam questions preparations. To evaluate the understanding of student in key course topic, lecturers can use a wide variety of actual examination formats including multiple-choice questions, true-false, fill-in-the-blank, short answer, problem-solving exercises, and essay questions [13]. This research focuses on generating multiple choice exam questions for two courses of the degree program which are Fundamentals of Computer Problem Solving and Introduction to Computer Programming. The remainder of the paper is structured as follows. We first present the literature review; this is followed by the methodology which describes the steps taken to develop our automated exam question generator prototype. Section IV is devoted to results and discussions and section V concludes the paper.

II. LITERATURE REVIEW

In this section we describe the six levels of Bloom's Taxonomy cognitive domain and brief overviews of the many techniques that have been used to generate the automated exam question generator.

A. Bloom's Taxonomy

According to Demetriulias [14], if students are to transfer knowledge into practice, testing is a must to foster the student's ability to think quickly and analytically, rather than only to recall facts, In 1956, Benjamin Bloom headed a group of educational psychologists who developed a classification of levels of intellectual behavior important in learning. During 1990's new group of cognitive psychologists, led by Lorin Anderson (a former student of Bloom), updated the taxonomy to reflect its relevancy to 21st century work. Bloom's taxonomy is a classification system of educational objectives based on the level of student understanding necessary for achievement or mastery. Educational researcher Benjamin

Bloom and colleagues have suggested six different cognitive stages in learning; from the simple recall or recognition of facts, as the lowest level, into the increasingly more complex and abstract mental levels, to the highest order which is classified as evaluation [15].

In order to effectively assess students' performance based on learning outcome, the exam questions should cover all six levels of Bloom's taxonomy. Bloom's Taxonomy is an educational objective categorization system based on the students' competency level and it emphasizes on the need to differentiate learners based on varied competency skills set [1] [3]. The standard and quality exam questions depend on a varied set of parameters including different level of learners and course objectives [1].

B. Current System of Automated Exam Questions Generator

Nowadays, there are many researches being conducted on automated exam questions generator. Automated exam questions generator can be categorized based on the algorithms that are utilized in order to generate the questions. Generally these generators can be divided into three types of algorithm such as random algorithm, backtracking algorithm and artificial intelligence algorithm [4]. Random algorithm and backtracking algorithm is widely used because of its simplicity [4]. Unfortunately, random algorithm tends to produce invalid exam question because it extract questions randomly based on exam paper constraint conditions [4]. Even though the backtracking algorithm can reduce the probability of invalid exam questions, but it faces the same problem as the random algorithm which is time consuming to search or backtracking when the questions volumes is huge [4]. Therefore, researchers started to adopt artificial intelligence in their researches to improve on the performance of Automated Exam Questions Generator and the quality of exam questions. Examples of the artificial intelligence algorithm are ant colony algorithm, simulated annealing algorithm and Genetic Algorithm [4].

Compared to other methods, Genetic Algorithm method has few advantages. First, it is fault tolerance; which means that it can progressively improve false or non-optimal solution towards an optimal solution [16]. Other than that, when compared to other exam questions generating methods such as Shuffling algorithm and utility-based agent, Genetic Algorithm has higher processing speed, generation success and quality ratio [6]. When compared to random method, Genetic Algorithm can produce a better composition of question [17] and conform to constraint conditions which lead to better rationality and easier control of difficulty coefficient [18]. Backtracking algorithm, while better than random, is surpassed by Genetic Algorithm in term of processing speed when processing large number of test papers [18].

Automatic Question Paper Generator System (QGS) [5] is a prototype that combined web-based and desktop-based with the main feature is to produce a non-duplicated set of exam paper. This system used the shuffling algorithm for randomization. The disadvantage of shuffling algorithm is it only focuses on generating a random selection of exam paper without any duplication and do not cover the difficulty and level of exam question [6].

Automatic Test Paper Generation Based on Ant Colony Algorithm [4] adopted binary ant colony algorithm in searching and selecting test paper automatically; and target weighting method to model the automatic test paper generation. The main focus of Automatic Test Paper Generation Based on Ant Colony Algorithm is searching performance rather than the quality of test paper generated where the difficulty and level of test paper are ignored [4].

Test Paper Based on Simulated Annealing Algorithm [7] introduced simulated annealing algorithm to optimize test paper problem based on the standard Genetic Algorithm. It used six attributes, i.e. score, type, difficulty, knowledge, chapter and time in the calculation for generating test paper questions, but only general difficulty levels are used i.e. easy, medium and difficult [7].

The random search algorithm can be categorized as instance-based or model-based. The instance-based methods generate new candidate points based on the current point or population of points, e.g. simulated annealing algorithm, Genetic Algorithm and tabu search. Meanwhile, model-based methods rely on an explicit sampling distribution and update parameters of the probability distribution, e.g. ant colony optimization, stochastic gradient search and cross-entropy method [8].

Automatic Question Paper Generator System (QGS) [5] and Automatic Test Paper Generation Based on Ant Colony Algorithm [4] did not cover the difficulty and level of exam question. Test Paper Based on Simulated Annealing Algorithm [7] implemented chapter selection and general difficulty level i.e. easy, medium and difficult which not based on Bloom's Taxonomy. On the other hand, the prototype of auto-generator examination question use Genetic Algorithm [6] and adopted Bloom's Taxonomy but did not apply chapter selection.

The prototype of auto-generator examination question using Genetic Algorithm [6] applied two main techniques; text matching for matching keywords of cognitive domain in Bloom's Taxonomy with the questions for sorting purpose and Genetic Algorithm to process the sorted questions into new set of questions' combination. The fitness function used is based on user input which includes number of questions needed, percentage of each of the taxonomy level and total score for the paper. Crossover method used was single crossover with random crossover point. For mutation, a question from target chromosome will be taken out and replaced with random question from question bank. Unfortunately, this prototype does not include chapters of the course syllabus into consideration and tends to generate questions which are biased on the course syllabus [6].

The purpose of this research is to develop a prototype of Automated Exam Questions Generator that is able to generate Exam Questions based on Bloom's Taxonomy level and chapters chosen without repeating previous questions in two consecutive years. This research focuses on instance-based random search algorithm and specifically focuses on Genetic Algorithm because Genetic Algorithm can realize global optimization with a fast rate of convergence [4]. Genetic Algorithm is an adaptive heuristic search based on evolutionary of natural selection and genetics; and it also exploit random

search and have objective or lead to the result of the search to optimize the selection [6]. More explanation on Genetic Algorithm will be presented in Section III.

III. METHODOLOGY

We divided this section into three phases where we will describe the theoretical framework in the first phase, the development of the Automated Exam Question Generator in the second phase and the Genetic Algorithm for the Automated Exam Question Generator in the final phase.

A. Theoretical Framework

The key factors for Automated Exam Question Generator are:

- The exam questions generated will be higher quality if the question bank contains all questions from all chapters for the subject and covers all six levels of Bloom's Taxonomy.
- The better categorization of exam questions based on Bloom's taxonomy stored into questions bank, the high quality of exam questions can be produced.
- The exam questions selections are varied and the chances to repeat the same question are less; if the size of the question bank is enormous.
- The better algorithm created, the faster quality exam questions can be generated.

The theoretical framework for Automated Exam Questions Generator as shown in Figure 1 below:

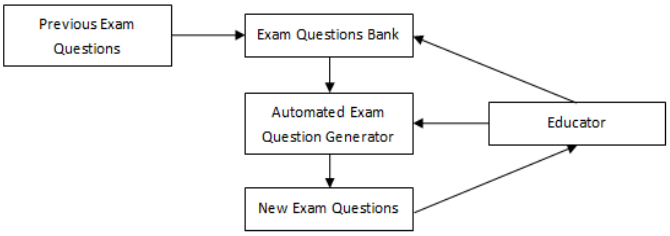

Fig. 1. Theoretical Framework.

Previous exam questions will be processed and stored in questions bank for the input for Automated Exam Questions Generator. The educator may also create new exam question and store into question bank for future use. The educator can choose the chapters to be covered in exam questions. New exam questions will be generated once educator trigger generate button. The Automated Exam Questions Generator will generate new exam questions based on six levels of Bloom's Taxonomy and chapters selected by the educator.

B. Automated Exam Question Generator

During the system development, all procedures and data requirements must be analyzed and documented. The interface design of Automated Exam Question Generator is shown in Figure 2 below:

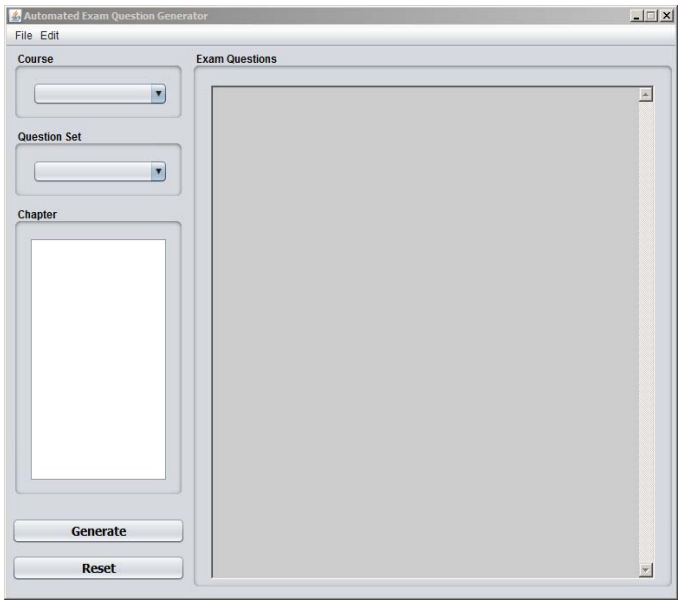

Fig. 2. Automated Exam Question Generator Interface Design.

The software requirements for Automated Exam Question Generator are shown Table 1:

TABLE I. SOFTWARE REQUIREMENT

Type	Description
Programming Language	Java version 1.8.0_65
IDE	NetBeans 8.1
Database	MySQL version 5.6.16

We will utilize Java as the programming language to develop the prototype, NetBeans will be used as an Integrated Development Environment (also known as IDE) and MySQL will be used for structural database to develop Automated Exam Question Generator. The architectural design of Automated Exam Question Generator is described in the following Figure 3:

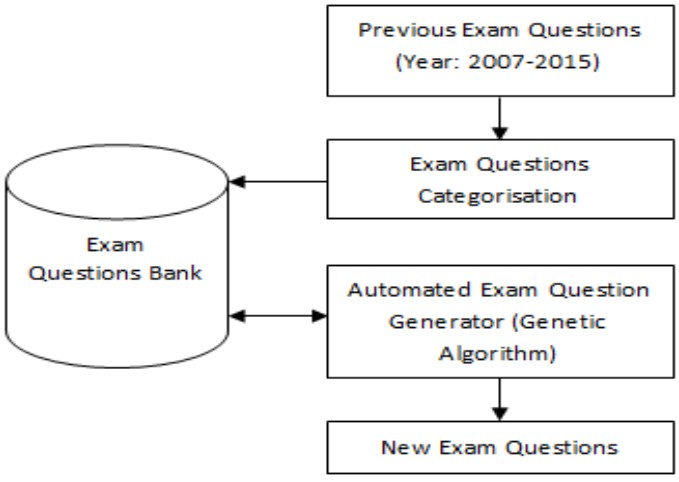

Fig. 3. Automated Exam Question Generator Architectural Design.

The previous exam questions from year 2007 until the year 2015 will be extracted and stored with its Bloom's Taxonomy level in Exam Questions Bank using a simple word matching function where Bloom's Taxonomy Level for each question from previous exam questions can be categorized based on the word in the exam question with Bloom's Taxonomy word list [11]. After word matching function is executed, the uncategorized previous exam questions; where there is no match found for the question and Bloom's Taxonomy word list will be updated manually. The Automated Exam Question Generator will generate new exam questions based on Bloom's Taxonomy Level and chapters chosen using Genetic Algorithm. The Entity Relationship Diagram (ERD) of Automated Exam Question Generator database design as display below in Figure 4:

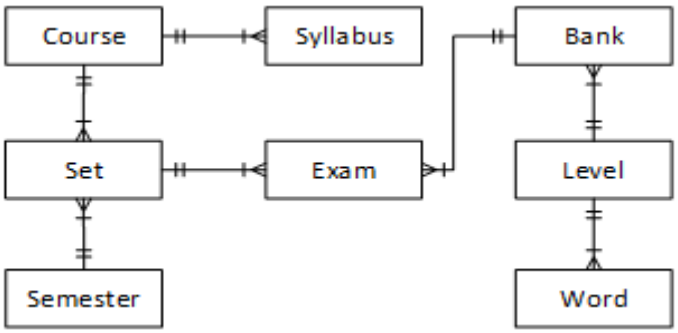

Fig. 4. ERD of Automated Exam Question Generator.

Here are the tables which will store the related details:

- **Course:** Course related e.g. course name and code.
- **Syllabus:** Syllabus related e.g. chapter and chapter title.
- **Bank:** Previous exam questions.
- **Level:** Six levels of Bloom's Taxonomy.
- **Word:** Word list of Bloom's Taxonomy.
- **Semester:** Semester related e.g. semester code, month and year.
- **Set:** Exam questions set details e.g. semester code and course code.
- **Exam:** Exam paper details e.g. exam set and question's code in exam questions bank.

C. *Genetic Algorithm*

The generic process of Genetic Algorithm [9] [10] is listed below:

- Step 1: Initialization
- Step 2: Evaluation
- Step 3: Selection
- Step 4: Crossover
- Step 5: Mutation
- Step 6: Repeat step 2 until a desired solution is obtained

Genetic Algorithm flowchart is shown in Figure 5 below:

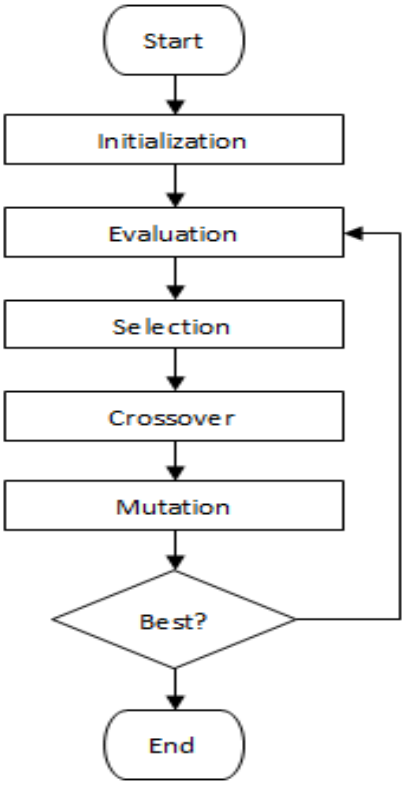

Fig. 5. Genetic Algorithm flowchart (Adapted from [9]).

Firstly, five initial populations will be created by randomly selecting questions from filtered questions bank. A simple query has been executed beforehand to filter all the questions in the bank so that the generation of new exam questions will consists only questions from chosen chapters and exclude previous questions from two consecutive years to prevent predictive questions pattern. Each population consists of collection of chromosome and each chromosome consists of genes. Each chromosome represent exam set and each gene represent a question in the exam set.

Secondly, each chromosome of the population will be evaluated and fitness value will be calculated for that individual chromosome. The fitness value will be calculated based on Bloom's Taxonomy classification (The six levels of cognitive domains with different degree of competency) and quality of exam questions weightage percentage as shown in Table 2 below:

TABLE II. BLOOM'S TAXONOMY CLASSIFICATION

Level 1	Knowledge	Easy
Level 2	Comprehension	
Level 3	Application	Medium
Level 4	Analysis	
Level 5	Synthesis	Hard
Level 6	Evaluation	

The quality of exam questions weightage percentage in Table 3 below are determined based on the coverage of

Bloom's taxonomy classification where Knowledge and Comprehension levels are grouped as Easy; Application and Analysis levels are grouped as Medium; and Synthesis and Evaluation levels are grouped as Hard.

TABLE III. QUALITY OF EXAM QUESTIONS WEIGHTAGE PERCENTAGE

Good	6 Level	e.g. Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation	100%
	5 Level	e.g. Knowledge, Comprehension, Application, Analysis and Evaluation	90%
	4 Level	e.g. Knowledge, Comprehension, Analysis and Synthesis	80%
	3 Level	e.g. Comprehension, Analysis and Synthesis	70%
Medium	4 Level	e.g. Comprehension, Application, Analysis and Evaluation	60%
	3 Level	e.g. Knowledge, Application and Evaluation	50%
	2 Level	e.g. Analysis and Synthesis	40%
Bad	2 Level	e.g. Application and Analysis	30%
	1 Level	e.g. Synthesis	20%

The good quality of exam questions covers all three Bloom's taxonomy classification i.e. Easy, Medium and Hard. The medium quality of exam questions consist at least two Bloom's taxonomy classification; either Easy and Medium combination or Medium and Hard combination or Easy and Hard combination. The bad quality of exam questions only covers one Bloom's taxonomy classification.

The formula in (1) is used to calculate the fitness value for each chromosome. W is the value of exam questions quality weightage percentage shown in Table 3.

$$1 / W \quad (1)$$

Thirdly, the best population is indicated by the lowest value of fitness function amongst five initial population and the selection is done using Roulette Wheel selection [6][9]. Fourthly, single crossover technique is used with the crossover point chosen randomly to enhance the fitness value of generated chromosome [6]. The example is shown in Figure 6.

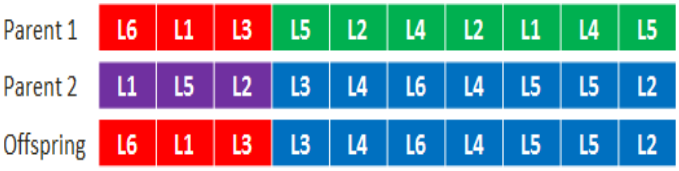

Fig. 6. Example of single crossover technique that produces an offspring chromosome with each gene contains Bloom's taxonomy level.

Fifthly, the mutation process will replace few genes with new genes based on mutation rate. In other words, questions from exam will be take out and replace with new selected questions based on given mutation rate. The example is shown in Figure 7.

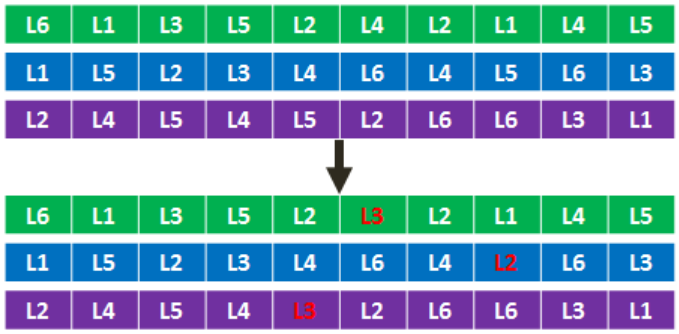

Fig. 7. Example of mutation process that produces new genes contains Bloom's taxonomy level.

Lastly, step 2 (evaluation) until step 5 (mutation) will be repeated until desired solution or maximum loop is obtained. All fitness value will be compared to find the lowest fitness. Optimal fitness is evaluated based on coverage of Bloom's Taxonomy level in generated exam questions.

IV. RESULTS AND DISCUSSION

The degree program usually consists of ten Multiple Choice questions, thus the total number of exam questions is initialise as ten. The Automated Exam Question Generator requires user to choose course code, exam question set and chapters before exam questions can be generated. The quality of exam questions weightage percentage is based on the level of cognitive domain covered and it has been defined as fitness value for this prototype. The low fitness value means that high quality exam questions will be generated. The Automated Exam Question Generator which consists of 500 sample questions in question bank; has been run 50 times to capture the fitness value of each test case to analyse the quality of exam questions generated by Automated Exam Question generator. Each test case using different number of chapters selected. The generated exam questions from this prototype shown in Figure 8.

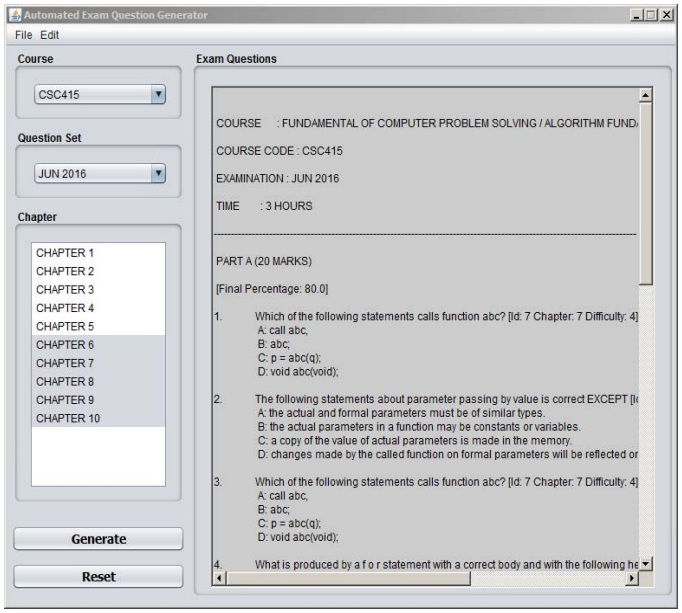

Fig. 8. Generated exam questions.

From the experiment conducted, the average value of exam questions weightage percentage is 70%. The highest exam questions weightage percentage is 90% and the lowest exam questions weightage percentage is 40%. The end result of this research affected by the smaller number of existing questions for each Bloom's Taxonomy level in question bank. The graph analysis of exam questions quality weightage percentage value for every test case is shown in Figure 9 below:

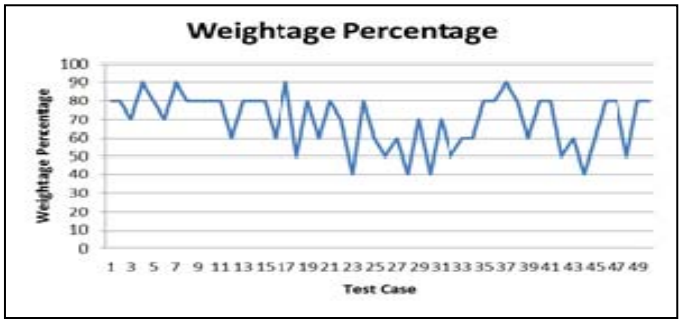

Fig. 9. Graph analysis of exam questions quality weightage percentage.

V. CONCLUSION

The Automated Exam Question Generator has been developed using Genetic Algorithm to reduce educators' burdens in the process of exam questions preparation. Currently, it focuses on Multiple Choice questions only. This research can be extended to use other types of questions such as short questions, essay questions, fill in the blank questions and true false questions. Furthermore it also can be used to generate quiz or test questions. In future, other type of algorithms can be applied to Automated Exam Questions Generator that has been developed in order to increase its performance and the quality of exam question generated.

ACKNOWLEDGMENT

This work is supported by ARAS Grant of Universiti Teknologi MARA (UiTM), Malaysia.

REFERENCES

[1] A. Immanuel and Tilasi.B, "Framework for Automatic Examination Paper Generation System," International Journal of Computer Science and Technology, vol. 6, no. 1, pp. 128-130, 2015.

[2] S. Pandey and K. Rajeswari, "Automatic Question Generation Using Software Agents for Technical Institutions," International Journal of Advanced Computer Research, vol. 3, no. 13, pp. 307-311, 2013.

[3] K. M. Boon and L. T. Lim, "An Examination Question Paper Preparation System with Content-Style Separation and Bloom's Taxonomy Categorisation," in The Third International Conference on E-Learning and E-Technologies in Education (ICEEE2014), Kuala Lumpur, 2014.

[4] D. Liu, J. Wang and L. Zheng, "Automatic Test Paper Generation Based on Ant Colony Algorithm," Journal of Software, vol. 8, no. 10, 2013.

[5] K. Naik, S. Sule, S. Jadhav and S. Pandey, "Automatic Question Paper Generation System using Randomization Algorithm," International Journal of Engineeting and Technical Research (IJETR), vol. 2, no. 12, pp. 192-194, 2014.

[6] N. H. I. Teo, N. A. Bakar and M. R. A. Rashid, "Representing Examination Question Knowledge into Genetic Algorithm," in IEEE Global Engineering Education Conference (EDUCON), Istanbul, 2014.

[7] J. Yang, "Design on Generating Test Paper Based on Simulated Annealing Algorithm," in International Conference on Civil, Materials and Environemental Sciences (CMES 2015), 2015.

[8] Z. B. Zabinsky, "IndE 516 Applications of Optimization in Engineering Design," 5 4 2009. [Online]. Available: http://courses.washington.edu/inde510/516/AdapRandomSearch4.05.2009.pdf. [Accessed 15 12 2015].

[9] D. Hermawanto, "Cornell University Library," 8 2013. [Online]. Available: http://arxiv.org/ftp/arxiv/papers/1308/1308.4675.pdf. [Accessed 22 10 2015].

[10] L. Jacobson, "Creating," 12 2 2012. [Online]. Available: http://www.theprojectspot.com/tutorial-post/creating-a-genetic-algorithm-for-beginners/3. [Accessed 18 12 2015].

[11] "Bloom's Taxonomy Verb List," University of Minnesota Duluth, [Online]. Available: www.d.umn.edu/vcaa/assessment/documents/Bloomsverblist.doc. [Accessed 10 12 2015].

[12] "ELECTRONIC QUESTION PAPER SYSTEM [PTAR UiTM]," UiTM, [Online]. Available: http://eqps.ptar.uitm.edu.my:8080/eqps/index.php?page=treeview/list&fac=CS. [Accessed 23 11 2015].

[13] Simkim, M. &. (2005). Multiple Choice Test And Student Understanding:What is the connection? Decision Science Journal of Innovative Education.

[14] Demetriulias, M. M. (1982). Constructing Test Questions for Higher Level Thinking. Lippicott-Raven.

[15] Bloom, B. E. (1956). Taxonomy of educational of educational objectives.The classification of educational goals:Handbook I.Cognitive domain. New York.

[16] L. Han and X. Li, "The Analysis of Exam Paper Component Based on Genetic Algorithm," 2014 Fourth International Conference on Communication Systems and Network Technologies, Bhopal, 2014, pp. 561-564.

[17] Rasim, A. Z. R. Langi, Y. Rosmansyah and Munir, "Generation quiz with Genetic Algorithm based on bloom's taxonomy classification in serious game based virtual environments," 2016 International Conference on ICT For Smart Society (ICISS), Surabaya, 2016, pp. 42-48.

[18] Kui Zhang and Lingchen Zhu, "Application of improved Genetic Algorithm in automatic test paper generation," 2015 Chinese Automation Congress (CAC), Wuhan, 2015, pp. 495-499.