Proposing an algorithm for selecting and evaluating the level of questions for Gifted Schools

Marwa Malik Hameed, [1] Prof. Dr. Dahir Abbas Reda, [2]

Iraqi Ministry of Education[1]
[2]Department of Information Technology, Technical college of Management, Iraq

Abstract : Generating tests from question banks using the elements extracted manually and randomly consumes a great deal of time and effort. The quality of the resulting tests is often inaccurate. The tests that have been created may not fully correspond to the requirements that were previously formulated, so this study focused on innovative ways to enhance this. The process by optimizing execution time and generating results closely meeting the requirements for extraction, this paper proposes a Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) in configuring the Hybrid PSP-GA algorithm to generate multiple-choice tests based on the difficulty levels of the questions and the extracted evaluations, shown Experimental results that PSO speeds up the extraction process, improves the quality of tests, and is more efficient in most criteria such as execution time, search area and contrast, and that the experiments and analyses with the proposed mutation operator have proven the success of the GA method with a very high rate, the research is distinguished by a hybrid PSO-GA technique to solve optimization problems, Further improving the balance between exploration and exploitation capabilities through the integration of genetic factors ,Designed an electronic system for selecting and evaluating For questions and their level of difficulty, based on the hybrid algorithm.

Keywords: Partial Swarm Optimization, Genetic algorithm, Hybrid PSOGA, Question Bank, MYSQL, Python, Web Application

Introduction

Education is an essential element in order to improve the country and its progress. Today, an important challenge facing education has reached a stage to enable universities and schools to provide more efficient, effective, and accurate educational processes .An important aspect of education is how to evaluate the questions placed for applicants to gifted schools in terms of quality, quality, level, and The goal many researchers have invested efforts to make computers work to automate the process of creating multiple-choice tests using available question banks and the results have been shown to be promising, and thus make multiple-choice tests more useful than the rest of the tests, and when you do not use any criteria to choose questions or use some Standards but it needs the participation of the academic to improve them If a lot of questions are identified, then human improvement will fail. There is a need for an algorithm to determine the most appropriate questions to complete the number of questions of specific test size and for the purpose of providing academics and teachers with effective means to conduct tests with multiple-choice questions from the question bank. The impetus for developing the PSO-GA approach is to combine the advantage of a genetic algorithm and improve particle swarm. By including genetic factors in the standard PSO, both models have strengths and weaknesses. In GA if an individual is not identified then the information contained in that individual will be lost, but the PSO has a fast memory however without a choice trigger. Thus the basic idea of PSO-GA is to combine the social thinking ability of PSO with the local search ability of GA because both GA and PSO are population-based algorithms, and the PSO-GA hybrid approach is also a population-based algorithm, thus Hurry up to find the best solution.

Research problem

The problem of the study lies in the lack of an automated system or a specific algorithm that selects and assesses the quality of the questions taken by the committee supervising the development of questions for students of gifted schools in the Iraqi Ministry of Education.

The goal of the research

Manual selection of questions from the database bank by supervisors has become something that is not suitable for our time due to the development of information technology, smart methods, and hybrid algorithms, so the study aims to build a hybrid algorithm for selecting and evaluating the level of questions for students of gifted schools in the Iraqi Ministry of Education from the question bank database.

– The Standard PSO algorithm

The PSO swarm particles algorithm was developed by Eberhart and Kennedy in 1995 and it simulates the social behavior of swarm particles and fish that these flocks use to find shelter, housing, food sources, or another suitable habitat. One of the intelligent technologies that are inspired by the evolving collective social behaviors of some organisms such as ants, bees, and mammals, and these algorithms have proven their effectiveness as they are in nature and have had a great role in the wonderful and effective applications that covered various fields. In nature, the PSO algorithm is a stochastic optimization technique that adapts well to optimizing nonlinear functions in multi-dimensional space and has been applied to many real-world problems [4], [5].

The particles of the swarm are stationed at random locations and move randomly within the search range, and the direction of the particles changes in search of new locations better than the previous locations. To find the best location for each particle of the swarm particles (BP) Best Position is done by finding the new velocity Vi (t + 1) and accordingly Relationship (1) and then the new position (BP) for each particle of the swarm is calculated according to a relationship (2). To find the best position for all the swarm particles, the Global Best Position (GP) requires the existence of a Fitness Function for each particle of the swarm and through it. We can know the global GB best position for all of the swarm particles , since the global best position corresponds to the highest value of the efficiency function[3] [6] [8].

$$V_{i+1} = V_i + C_1 R_1 (LB_i - x_i) + C_2 R_2 (GB - x_i) \ \dots\dots\dots \qquad (1)$$

Represents the velocity of the particleV_i

They represent randomly generated numbers within a range [0,1] R1,R2

C1,C2Weights constant

$$x_{i+1} = x_i + V_{i+1} \dots\dots \qquad (2)$$

The current position of the swarm particleX_{i+1}

X_iThe previous location of the swarm particle

V_{i+1} The new velocity

- The standard genetic algorithm GA

Genetic algorithms (GA) is a parallel and universal search technique that mimics natural genetic factors. GA applies factors inspired by the mechanisms of natural selection from binary strings. The genetic algorithm is one of the important methods of artificial intelligence, where the work of the genetic algorithm begins with a first set Of the chains that are generally called chromosomes and are randomly selected an elementary community, the chromosome is a representative solution to the problem and represents a group of gene and the size of the chromosome varies from one problem to another (population size), then the probability of each chromosome remaining in the community is determined in other words, fitness, and fitness. An evaluation mechanism for the chromosome, whereby the selection of candidate individuals from the candidate chromosomes in the next generation is based on fitness values, this means choosing only the genetically new individuals (chromosomes) to be parents, after selecting the cross and mutation respectively. Crossover and Mutation and a random point are chosen for the crossover, after which the candidate chromosomes undergo a mutation. update on bits or genes. The mutation process expands the search area to areas that may not be close to the current chromosome community, thus ensuring a holistic search. This evolution cycle is repeated until the required standard is reached. This criterion can be determined by the number of development cycles (arithmetic operations). Or a predefined value for the target function. [1], [2]

-Hybrid PSO-GA algorithm

The impetus for developing the PSO-GA approach is to combine the advantage of a genetic algorithm and improve a particle swarm. By including genetic factors in the standard PSO, the balance between exploration and exploitation capacity is further improved, both models have strengths and weaknesses. In GA, if an individual is not identified then the information that that individual contains will be lost, the basic idea in PSO-GA is to combine the social thinking ability of PSO with the local research ability of GA, Since both GA and PSO are population-based algorithms, and that the hybrid PSO-GA approach is also a population-based algorithm, thus speeding up the global solution. The approach starts from the initialization stage, in which the particles of the swarm and their corresponding velocities are randomly generated over the search area, thus, the position of each particle in the swarm is modified according to its own experience and that of its neighbors. The former is known as the best personal website (pbest) and the latter is the gbest placement. After iteration, the position of each particle is updated after the formation of a new generation in the PSO iterations, some particles of the new community are selected and then GA is applied. The algorithm aims to create a new population by replacing points in the current group with better points via genetic principles, that is, by performing selection, crossover, and mutation operations. Then Gbestis returned with the best set of questions. [5], [7]

Marwa Malik Hameed, [1] Prof. Dr. Dahir Abbas Reda, [2]

-Proposed system of PSO-GA Hybrid algorithm
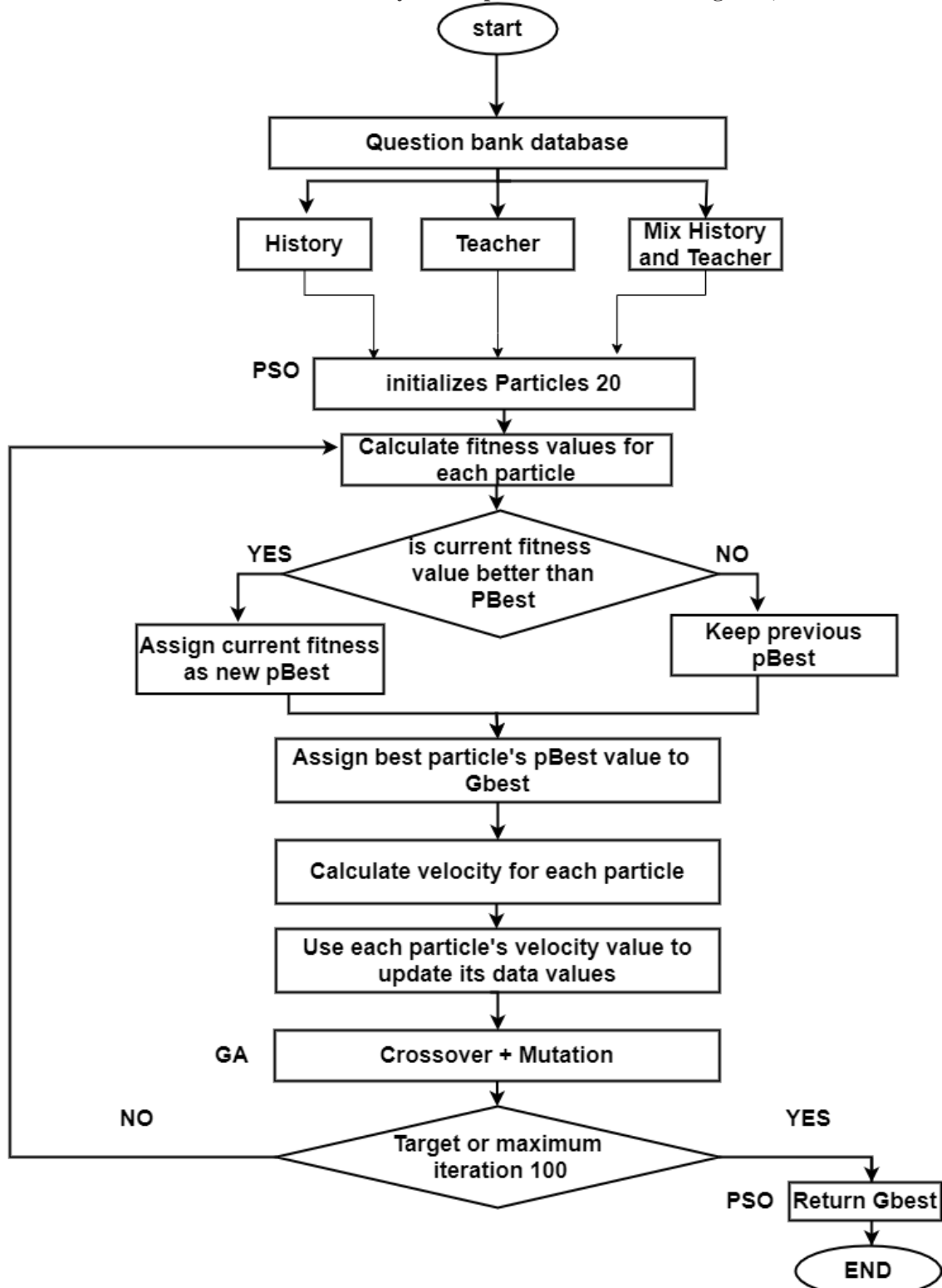**The system steps are summarized in Figure 1, which includes: -**

Figure (1) includes the steps of the PSO-GA Hybrid algorithm

Figure (1) shows the question bank database, which is formed at this stage from three databases in MYSQL program. Each rule contains Subjects questions (intelligence, mathematics, science) and   consists each database as in Table (1) consists of the column (A) which represents the question number (ID) and column (B) the number of correct answers depending on the students 'answers, and the column (C) represents the total

number of answers to the question and the column (D) represents the efficiency ratio of each question based on the students' previous answers andthe column (E) represents the difficulty assessment the question is based on the professor's evaluation of the subject, andfinally Column (F) represents a common table for evaluating the question among the total percentages.

Table (1) clarification of the types and methods of evaluation of the questions

A	B	C	D=1-B/C		E	F=D+E/2
1	5	7	0.2		0.1	(0.2+0.1)/2=0.15
2	3	4	0.25		0.4	(0.25+0.4)/2=0.32
100	5	10	0.5		0.3	(0.5+0.3)/2=0.4

As for the selection of questions, it is through the History Fitness section, in which there is a question's difficulty rate depending on the students' answers, As for the professor's assessment section, questions are withdrawn depending on the column that contains the of their difficulty from the professor's of the subject, and it may be placed in the questions bank first time and finally, the Mix section in the second level of Figure (1) represents a method for determining the percentage of difficulty of the questions based on the students' answers and the assessment of the professor together, and Figure (2) shows the types and methods of selecting the questions.
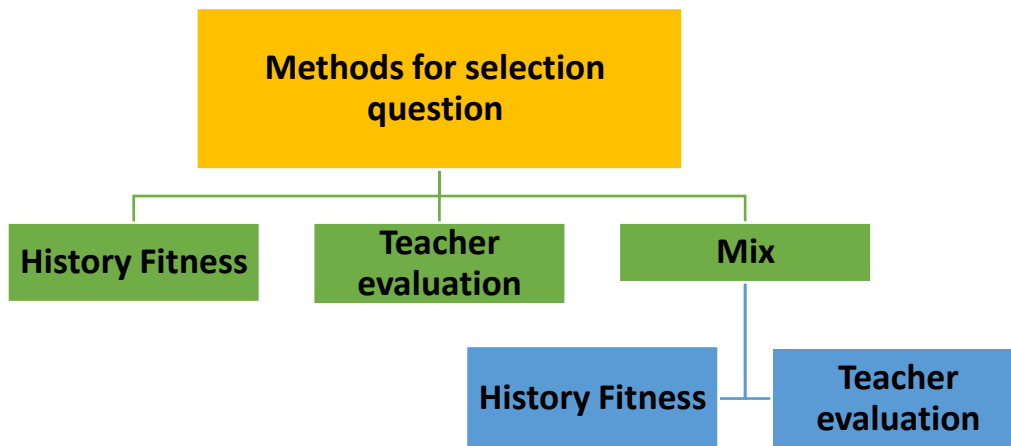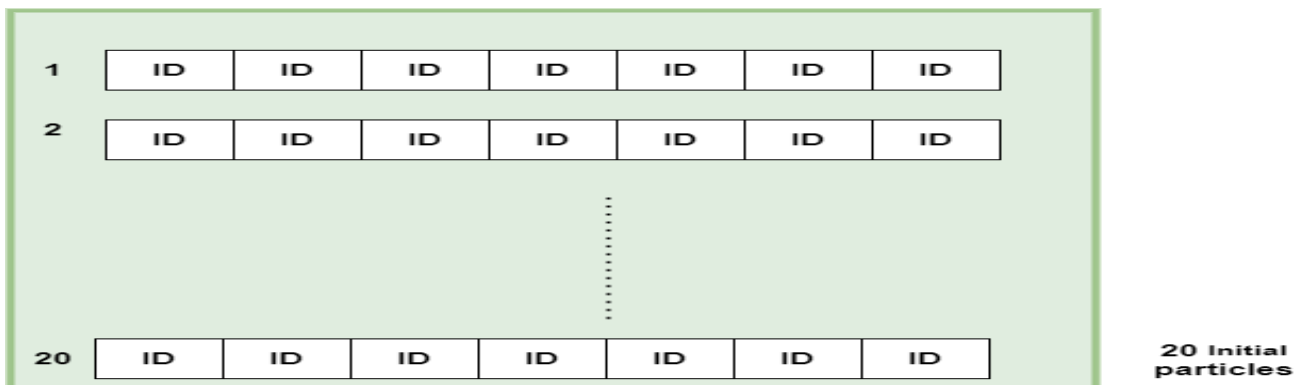

Figure (2) illustrates the types and methods of selecting questions

The third level of Fig. 1 represents the creation of a new question community using the PSO algorithm Which are generated randomly (initial particles) by creating the elements particles, as in Figure (3), which shows the number of elements chosen in the elementary community.

Figure 3 illustrates the creation of the initial particle community
Also the fitness of each question in the particle in the initial population is calculated by the relationship (1)

$$Fitness\ 1\ =\ 1 - \frac{Correct\ Ancwers}{ALL\ Ancwers}$$ **(The fitness for first cell in** particle**)**

$$Fitness\ 2\ =\ 1 - \frac{Correct\ Ancwers}{ALL\ Ancwers}$$

(1) **(The fitness for second cell** in particle**)**

$$Fitness\ 7\ =\ 1 - \frac{Correct\ Ancwers}{ALL\ Ancwers}$$

(The fitness for seven cell in particle**)**

The evaluation of the third selection method for questions can be calculated through the relationship (2) based on the percentage of students' answers and the ratio set by the professor divided by 2.

$$\text{Fitness}\ MIXQ_n = \frac{\text{History Fitness}Q_n + \text{Teacher evaluation}Q_n}{2}$$ …. (2)

The fitness of all questions in a single particle is measured through relationship (3).

$$F_{all} = (\ F1 + F2 + F3 \ldots\ldots + F7)\ /\ 7 \qquad \ldots\ldots \qquad (3)$$

F1 : The fitness of the first question in the particale
F2 : The fitness of the second question in the particale

⋮ F7 : The fitness of the last question in the particale (Identifying seven questions, for example)
The velocity can be calculated for each value (the number of selected questions, for example, 7 within a single particle in the initial population through the relationship (1)
Then calculate the new position of the values in the particle as in relation (2)

Movement and replacement process The content of the questions with some other questions is done according to the new candidate question number from the question bank, followed by the application of the GA genetic algorithm. Two elements of the initial population are selected after updating the relationships (1) and (2).
As for the process of crossover, a new generation of questions is produced by merging the first four genes from the first chromosome with the last three genes from the second chromosome and vice versa. Figure (4) shows the process of crossing over.
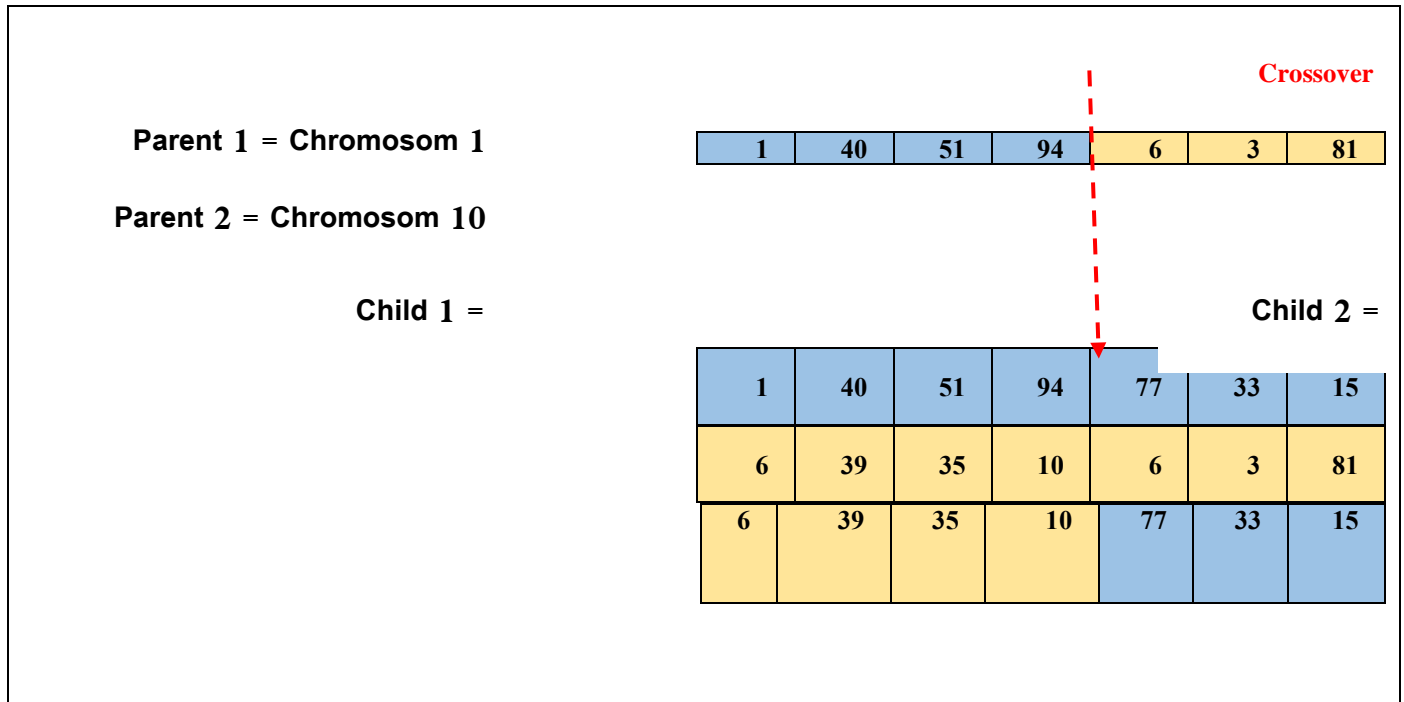
Figure (4) illustrates the process of intersection to produce a new generation of questions

Also, Figure (5) shows the process of mutation in one gene for each new chromosome, which must be unique to prevent a recurrence.
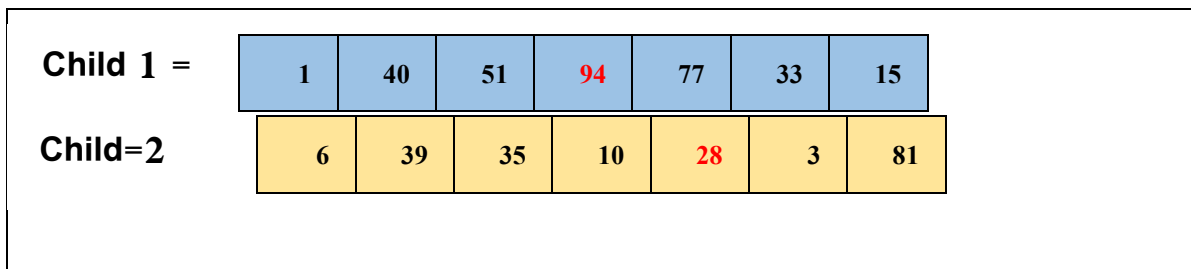

Figure (5) shows the process of mutation in one gene for each new chromosome

And finally, the update takes place in the initial population and then determines the best pbest and best gbest as described below

If Gbest<Pbest

Gbest =Pbest

The above operations are repeated as shown in figure (1) on each of the three rules (intelligence, mathematics, science) and according to the selection and determination of the number of questions to be selected from the base, to produce in the end a set of questions from the three rules Will be shown the Web Application for testing and evaluation**.**

The figure below (6) shows the work of the hybrid algorithm for the method of selecting the types of questions and extracting the desired questions.
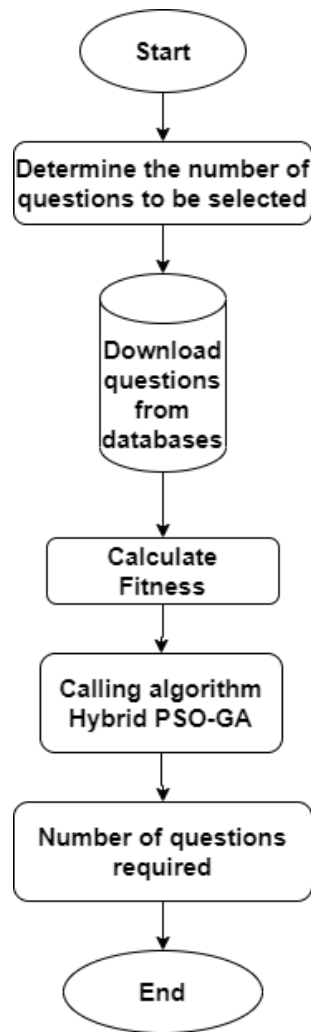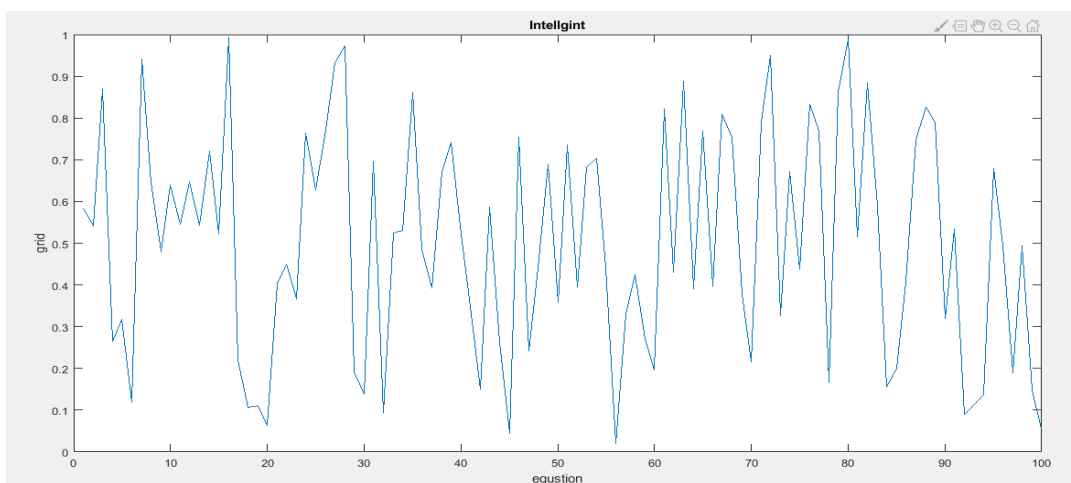
Figure (6) illustrates the steps for selecting questions in the hybrid algorithm

Figures (7, 8, and 9) show samples of the three rules that were created to show the difficulty rates of the questions for the three subjects .

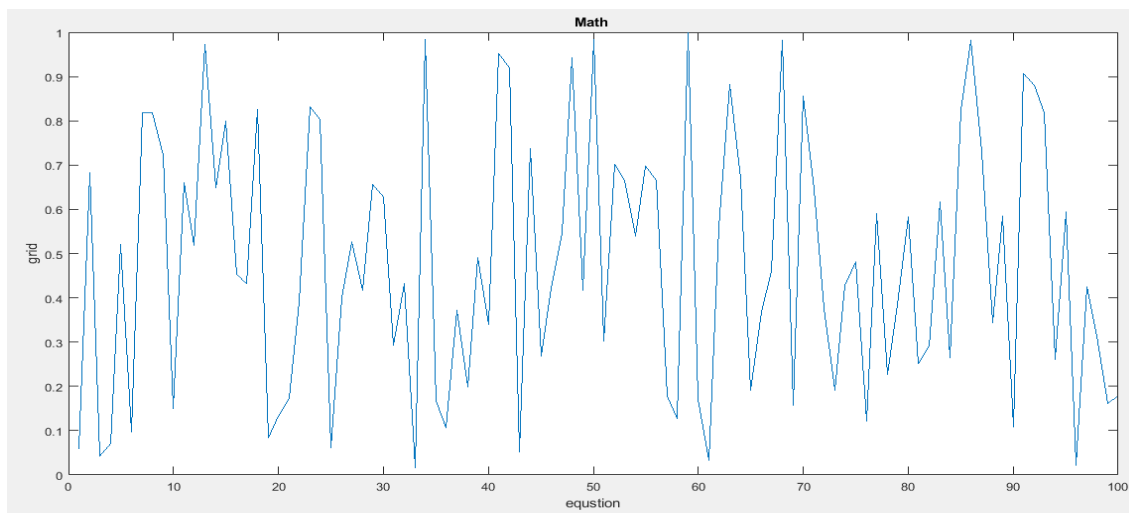Figure (7) the questions fitness percent for the intelligence subject

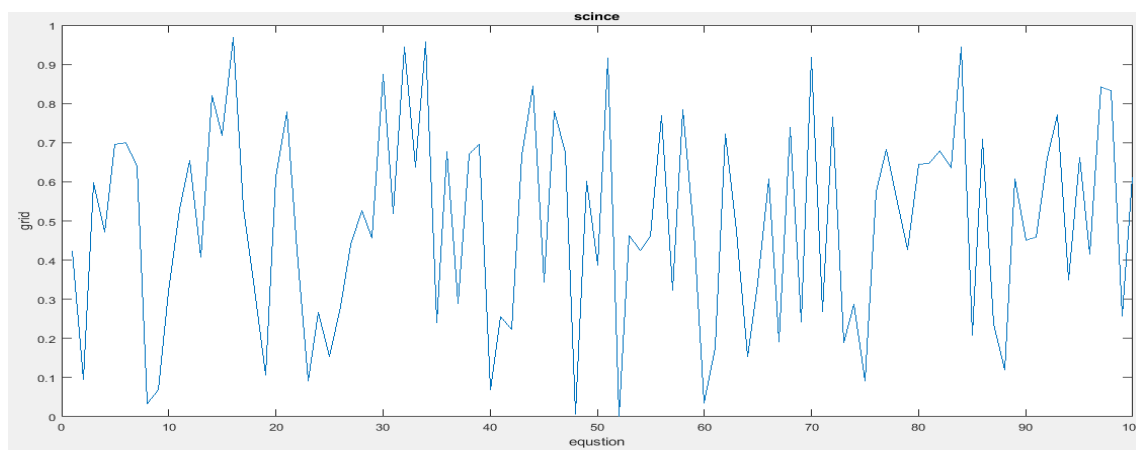Figure (8) the questions fitness percent for the math subject

Figure (9) the questions fitness percent for the sciences subject

-Data

The proposed system was applied to a database from the Gifted Schools Care Authority / Iraqi Ministry of Education that includes test questions for the years from 2010 to 2021 as well as the results of students 'answers to the questions, as the data and questions were collected from the above-mentioned schools after obtaining the official approval from the Iraqi Ministry of Education to use It consisted of three subjects: (intelligence - mathematics - science) and the percentage of each subject in the test was determined based on the opinion of teachers, and the design of the database included each table for one of the three subjects (the first for intelligence questions, the second for math questions, and the third for science questions) As shown in Figure (10)
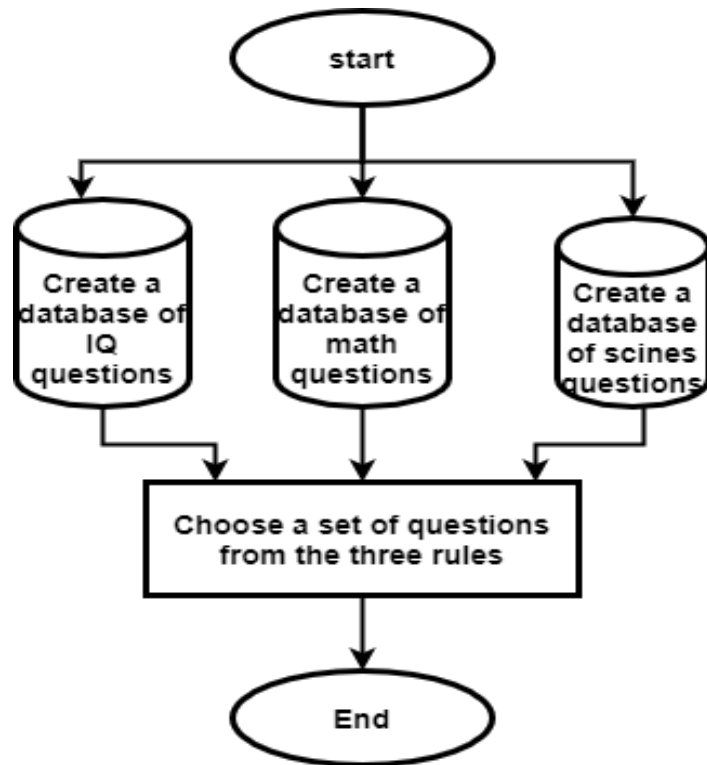
Figure (10) Description of the questions database

Method	Fitness	Questions Intelligence
History	0.7357859531772576	(qu5,qu15,qu9,qu99,qu65,qu77,qu98,qu34,qu8,qu25,qu20,qu22,qu68)
Teacher	0.6615384615384615	(qu75,qu1,qu3,qu80,qu46,qu33,qu58,qu39,qu38,qu55,qu21,qu28,qu6)
MIX	0.7692307692307693	(qu11,qu85,qu61,qu70,qu49,qu79,qu82,qu39,qu7,qu64,qu24,qu17,qu73)

Results -
Table (3) the results obtained from the application of the Hybrid PSO-GA algorithm for the math questions database

Method	Fitness	Questions Math
History	0.7792207792207791	(qu5,qu15,qu9,qu36,qu13,qu11,qu39)
Teacher	0.7714285714285715	(qu39,qu11,qu10,qu20,qu21,qu32,qu28,)
MIX	0.8	(qu4,qu37,qu16,qu21,qu12,qu38,qu25)

Table (4) the results obtained from the application of the Hybrid PSO-GA algorithm for the sciences questions database

As a result, 25 questions are extracted for the test after determining the method from which the questions are drawn for the three subjects.

Conclusions

The new hybrid gene particle swarm algorithm performs better than PSO and standard GA algorithms. The results showed that the hybrid algorithms take advantage of the advantages of both GA and PSO methods and are thus useful for better solution problem solving and offer faster convergence.

References
1. Yildirim, Mehmet. "A genetic algorithm for generating test from a question bank." Computer Applications in Engineering Education 18.2 (2010): 298-305.
2. Qasim Ibrahim, Yahya "Attacking an Encrypted Text (whose explicit text is known) using a genetic algorithm," Journal of Education and Science, 1.27 (2014): 158-168
3. Nguyen, Tram, et al. "Multi-Swarm Optimization for Extracting Multiple-Choice Tests From Question Banks." IEEE Access 9 (2021): 32131-32148.
4. Saleh, Ibrahim Ahmed, and SalhaRaaMahammead. "Apply Particle Swarm Optimization Algorithm to Measure the Software Quality." AL-Rafidain Journal of Computer Sciences and Mathematics 12.1 (2018): 26-36
5. Esmin, Ahmed Ali Abdalla, Germano Lambert-Torres, and GuilhermeBastosAlvarenga. "Hybrid evolutionary algorithm based on PSO and GA mutation." 2006 Sixth International Conference on Hybrid Intelligent Systems (HIS'06). IEEE, 2006
6. Khan, Koffka, and Ashok Sahai. "A comparison of BA, GA, PSO, BP and LM for training feed forward neural networks in e-learning context." International Journal of Intelligent Systems and Applications 4.7 (2012): 23.
7. Garg, Harish. "A hybrid PSO-GA algorithm for constrained optimization problems." Applied Mathematics and Computation 274 (2016): 292-30
8. Bui, Toan, et al. "Application of Particle Swarm Optimization to Create Multiple-Choice Tests." Journal of Information Science & Engineering 34.6 (2018).
9. [9] Borela, V. (2020). ALS Curriculum Implementation for Juvenile Delinquents in Urban and Rural Areas in the Philippines: A Comparative Analysis. Middle Eastern Journal of Research in Education and Social Sciences, 1(1), 1-14. https://doi.org/10.47631/mejress.v1i1.3
10. [10] Temsah, L., & Safa, N. (2021). New Approaches to Simulation-Based Science Instruction to Enhance Reasoning and Communication Skills in Lebanese Elementary Education. Middle Eastern Journal of Research in Education and Social Sciences, 2(1), 56-79. https://doi.org/10.47631/mejress.v2i1.180