Improved particle swarm optimization algorithm in professional skills identification management system

Yinghui Peng
Tianjin University of Technology and Education, Tianjin, China

Abstract. This paper adopt improved particle swarm algorithm to solve multi-objective optimization problem; so this project can determine vocational skills certification examination test of the different professions, different types and different levels of difficulty; the making test paper covers a wide range of knowledge and chapter score; and the difficult level of making test paper conform to normal distribution, meet different groups requirements of vocational skills certification examination.

1 Introduction
With the development of information technology and continuously improve the level of education, and educators intelligently put forward higher requirements for the speed and efficiency, after a test of test questions is not limited to a single subject, objective, requirements test questions are diverse, covering the wide range of knowledge, suitable degree of difficulty of the questions, More attention to test students ' knowledge and practical skills. In occupational skill testing questions management system, a single, slow group volumes long, remains a widespread problem, still cannot meet the needs of professional skill appraisal assessment. So I propose a realization policy of vocational skill identification bank management system based on improved Particle Swarm Optimization algorithm.

2 Theoretical study on occupational skill testing questions intelligently
Vocational skills appraisal test each system has a certain role, each of which has a certain meaning, each index parameter is for examination paper analysis, automatic test, examination management services, these parameters are optimized, system based on parameters also has some optimizations. Intelligent test paper composition parameters are as follows

2.1 Test scores M
Test scores $\sum_{i=1}^{M} K(i) a_{i2} = M$, $a_{i2}$ is a question I score, $K(i)(1 \leq i \leq M)$ as I question the state ,variable k(i) take the value 0 or 1.

2.2 Exam time T
The elapsed time to complete all the questions, most students finish and check the problem need consuming time, item writers can adjust it according to the experience in teaching comprehensive objective information; $\sum_{i=1}^{M} K(i) a_{i3} = T$, $a_{i3}$ is estimated time to finish the test I.

2.3 Chapter points
Course sections and the various types of questions in question bank exists relation, in order to facilitate assessment and management; $\sum_{i=1}^{M} K(i) a_{i4} a_{i4} C(j) = S_j$, $S_j$ is chapter j score, $a_{i4}$ is questions I belongs to section, $c(j)$ take the value 0 or 1.

2.4 Knowledge points score
Syllabus requires assessment of knowledge; $\sum_{i=1}^{M} K(i) a_{i5} O(j) = D_j$, $D_j$ is point j score, $a_{i5}$ is a question I owned knowledge, $O(j)$ take the value 0 or 1.Intelligent test paper composition of knowledge in selected to meet uniform and focus problems, through the creation of hierarchical tree structure of knowledge management chain, solve linear density of knowledge in the management of large or small problems, knowledge distribution, given in the hierarchy according to the weights of the weight calculation of two-level model knowledge point reference weights build classification knowledge points are evenly weighted selection algorithm.

2.5 Paper type fraction
Paper type change because of the different professions,
different types of user needs, usually set to 5 type: judgment questions, multiple choice, fill-in questions, short answer, essay questions, etc; \( \sum_{i=1}^{M} K(i)_{a_{ij}} T(j) = F_j \), \( F_j \) is score in j type, \( a_{ij} \) belong to the types of questions I, t (j) take the value 0 or 1.

### 2.6 Ability Level score

\[ \sum_{i=1}^{M} K(i)_{a_{ij}} A(j) = G_j \], \( G_j \) is j ability level score, \( a_{ij} \) questions I belong to ability level, a (j) take the value 0 or 1.

### 2.7 Test difficulty

Difficulty is a measure of test questions and test measure of the degree of difficulty. Difficulty is obtained by calculation of students' difficulty value of test. \( \sum_{i=1}^{M} K(i)_{a_{ij}} a_{ij} / M = DIF \), DIF is general degree of difficulty, \( a_{ij} \) is the difficulty coefficient of test I, Fuzzy five levels listed in the following table I.

### 2.8 The degree of examination questions

Questions of discrimination is an important indicator of performance, is a measure of the difference between test ability, difficulty and discrimination are closely linked discrimination of the questions is Max at the time of 0.3-0.6 difficulty. Method of calculating distinction is that test scores of the candidates are arranged from the highest score to the lowest score, and the highest scores group is from 27% of total scores, and the lowest scores group is from 27% of total scores. \( D = (S - S_L) / S \), \( S_H \) and \( S_L \) respectively are average scores of the highest and the lowest scores , \( S \) represents the full mark of the question, \( D \) represents the degree of examination questions, the greater of the \( D \) represent the greater of testers score differences. Difficulty evaluation criteria listed in the following table II. \( \sum_{i=1}^{M} K(i)_{a_{ij}} a_{ij} / M = DIS \), DIS is distinction of the papers, \( a_{ij} \) is the distinction of questions I.

The more difficult questions and the more easier questions should be fewer in proposition, the questions of the medium difficulty are more larger , Makes questions of average difficulty close to 0.5, difficulty and distinguish degrees is a relative indicators for groups, no absolute difficulty and distinguish degrees, if different capacity of test students groups are tested with same questions,difficulty and distinguish degrees of questions also will different, according to students of knowledge capacity level situation general rendering distribution of features, proposition teachers determine meaning of initial value, then further amended according to students' questions exam, and set the basis for next item difficulty and discrimination.

### 3 Occupational skill testing questions intelligently mathematical modeling

In vocational skills appraisal test this algorithm, the strategy of intelligent test paper Bank system data source is a matrix, as shown in equation 3.1.

\[
A_{m \times n} = \left( \begin{array}{cccc}
    a_{i1} & \cdots & a_{in} \\
    \vdots & \ddots & \vdots \\
    a_{m1} & \cdots & a_{mn}
\end{array} \right)
\]  

(3.1)

Where m is the total number of test questions, n is the number of attributes, \( N = 9 \), attribute indicators for each question that is the n-dimensional vectors, including scores, item type, degree of distinction etc. \( 8 \) indicators. All properties of each question is the rows in the matrix, as column respects each property of all the questions. Generate test constraints up to n−1, each column is question number in the destination matrix.

Generating test paper process is analysis test multiple constraint parameters based on the purpose of the test, searching rows \( x = (x_1, x_2, x_3, \ldots, x_p) \) of meeting the requirements in the matrix shown in 4.1.where \( p \) is the number of questions, \( X_i = (x_1, x_2, x_3, \ldots, x_{ij}) \) is vector composted of \( i \) type questions, \( X_{ij} \) is j question of I type.

### 3.1 Intelligent test paper composition function

Test paper problem is a multi-objective problem solving process, under normal circumstances, constraints above and generating test request error make objective function \( f \), when \( f \) minimum, generated the papers wanted by us, thus, multi-objective optimization problem transfer into single-objective optimization problem, the objective function such as

\[
\text{Min} f = \sum_{j=1}^{p} w_j x_j + \sum_{j=1}^{n-1} \sum_{i=1}^{m} a_{ij} x_{ij} + \sum_{j=1}^{n} (\sum_{i=1}^{m} a_{ij} x_{ij}) = \sum_{j=1}^{n} F_j
\]

\[
\sum_{j=1}^{n} (\sum_{i=1}^{m} K(i)_{a_{ij}} a_{ij} T(j) + d_{ij}) = G_j
\]

\[
\sum_{j=1}^{n} (\sum_{i=1}^{m} K(i)_{a_{ij}} a_{ij} T(j) + d_{ij}) = G_j
\]

(3.2)

In the optimization of the objective function above, \( d \) is deviation in the intelligent test paper generation between question properties and evaluation requirements , \( w_j (j = 1, 2, 3, \ldots, 8) \) represent weight of the target group, the sum of weight is to 1.

### 3.2 The steps of generating test paper

1. Determine subject knowledge points and scope of the examination, the examination content includes sections and scope. In the specified test range, clear analysis of knowledge content, focus, difficulties and a clear ability to assess students' ability level mastering of knowledge of each point.
2. Determine the distinction and difficulty of paper
(3) Determine type of test papers. Generally include blanks, choice questions, judgment questions, short answer and essay questions.
(4) Determine the number of test questions and test time
(5) Determine the paper structure, namely, chapters of the proportion of scores in the paper.
(6) Generating test paper, according to papers information, form a set of paper as possible to meet the requirements of the papers.
(7) Output, adjust the layout and output the final test paper.

3.3 Basic idea of particle swarm optimization algorithm

Particle Swarm algorithm is inspired by the bird foraging behavior which is a new type of swarm intelligence algorithm after Ant Colony algorithm. Particle Swarm Optimization algorithm description: total of n particles, the position of I particles is to X=(Xi1, X2,…, Xin), each particle has a measurable quality Xi's fitness; particle's velocity of the I-th particle (location change ratio) is expressed as Vi=(Vi1, V2,…,Vin)

Particles passing through the individual extreme value XIr=(Xi1*, Xi2*, …, Xin*) and a global optimization g*(g1,g2,g3….gD) to update the velocity and position. The speed and position of the particle would be updated according to the following equation

\[ V_{i,j+1} = V_{i,j} + \alpha r_1(x_{i,j} - x_{i,j}) + \beta r_2(g^* - x_{i,j}) \]  
\[ X_{i,j+1} = x_{i,j} + V_{i,j+1} \Delta t \]  

Vi is the previous speed, \( \alpha r_1(x_{i,j} - x_{i,j}) \) called cognitive part, \( \beta r_2(g^* - x_{i,j}) \) affect the next behavior of particles. Xi and VI are the position and velocity of particle I, the searching optimal position of particle I is Xi* and g* are searching optimal locations of current Particle Swarm optimal, w is inertia parameters and control the dynamics of particles flying, \( \alpha \) and \( \beta \) are called parameters or constant acceleration.

In General, access \( \alpha = \beta = 2.1 \) is the number of iterations, the R1 and R2 are random numbers in range of the \( [0, 1] \). Random initial position and velocity of particle swarm, then iteration using the above equation until you find the optimal solution.

3.4 Algorithm steps

(1) Initialize the particle position
(2) Number of iterations I=I+1
(3) Initial value of i=0 and fitness evaluation
(4) Calculation groups best historical position
(5) Update the particle swarm location according to location update formulas
(6) I=I+1
(7) Determine if I is greater than n, if established, go to step 8, otherwise, go to step 3
(8) Determine whether the termination condition is met, if established, end program after the output. Otherwise, go to step 2

3.5 Improved Particle Swarm Optimization algorithm

In order to accelerate convergence, Accelerate the Particle Swarm Optimization algorithm has been proposed by a doctor of Cambridge University in 2008, speed updated is done by a simpler formula as 3.5 below.

\[ V_{i,j+1} = V_{i,j} + \alpha r_1(t) + \beta r_2(g^* - X_{i,j}) \]  
\[ X_{i,j+1} = X_{i,j} + V_{i,j+1} \Delta t \]  

Where r is a random number and can be obtained by normal distribution, in essence, updating formula is the same as formula (3.3), using formula (3.5) instead of the speed in formula (3.6), you can obtain a formula which can represent location update status, such as formula (3.7) shows.

\[ X_{i,j+1} = (1-\beta)X_{i,j} + \beta g^* \]  

Speed does not appear In equation (4.7) speed may not be updated in improved Particle Swarm algorithms, therefore the algorithm simpler and easier to implement. R is obtained by a probability distribution, and make the algorithm does not fall into any local optimal solutions. Under normal circumstances, \( \alpha = 0.7-0.9, \beta = 0.2-0.7. \)

Under normal circumstances, standard acceleration particle swarm algorithm has strong ability to search exploration, and can be find the area of the best solution, but its mining solution is poor weak and cannot always guarantee convergence to the optimal solution. Now add a variation factors in the original acceleration Particle Swarm Optimization algorithm, so can greatly improve the speed of convergence of the algorithm, and solve more practical problems. The introduction of variation factors can accelerate the PSO, to increase the diversity of the population, improve search efficiency and accelerate convergence of optimal solutions,

\[ X_{i,j+1} = x_{i,j} + F^*(x_{i,j} - x_{i,j}^*) \]  

Where f is the differential evolution variants, r1,r2,r3 is uniformly distributed random integer between 1 and N, N is the number of population. Using micro differentiation of variants to fine-tune the new solutions in order to increase diversity and improve search efficiency.

3.6 weight determination of the intelligent test paper

Weight is a number ranging from 0 to 1, and all top level indicator weights sum to 1, weight is scientific and oriented, weight of important need to increase, and weight of not important indicators need small. Scientific distribution of weight indicator is also an important step in composing test paper.
According to the required 8 constraints in the system to make the test object weight \( W_j \) \((j=1, 2, 3, ..., 8)\), through the analysis of 8 constraints will break it into a tree structure, designed a NXN square matrix according to the constraints, while \( A^t \lambda I = 0 \), nonzero vector can be the eigenvectors of \( a \), \( \lambda \) is the eigenvalues of the matrix \( A \).

Get a relative value of \( a_{ij} \), forming a square matrix \( A \), suppose there is a normalized eigenvector \( A \omega = \lambda_{\text{max}} \omega \), \( \lambda_{\text{max}} \) is the largest eigenvalue of square matrix \( A \), \( \omega \) is the normalized eigenvectors corresponding to the \( \lambda_{\text{max}} \). \( \omega \) is the weights factors. If matrix satisfy three conditions

1. \( a_{ij} = 1 \)
2. \( a_{ij} = 1/a_{ij} \)
3. \( a_{ij} = a_{ik}/a_{jk} \), exactly \( \omega \) is the weights factors.

### 3.7 Intelligent generating test steps based on improved Particle Swarm algorithm

1. According to reorder the questions to test bank questions, questions can be divided into different subsets, and then renumber all the questions
2. According to the questions, calculate the number of questions to get total number of test papers
3. Computational Intelligence group volume weight of the index

(4) Initialize the Particle Swarm, random initialization of particle swarm, each particle \( X_i \) in particle swarm represents a possible group policy. In a population, the status of each particle adopt an \( m \)-dimensional vector \( X = (X_1, X_2, ..., X_m) \); the dimension of the vector \( X \) depend on the total number of questions; \( X_i \) is the components of the vector \( X \) and \( X_i \) \((i=1, 2, ..., m)\) is question number, question number cannot be repeated, according to the different type questions, number of examination questions should be arranged according to different segments

5. Calculate the fitness of particles according to the formula (4.2), and obtain the global extreme of current population

6. Save the best particle

(7) Particle update: update each particle location in the population according to formula (4.3, 4.4)

(8) Update optimal location \( g^* \)

---

![Program Flow Chart](image-url)
4 Introduction to occupational skill testing question management system

Occupational skill testing question bank management system functions should be developed as follows:

4.1 Account management module

Contain users of information, a variety of roles and system permissions management.

4.2 System settings module

Administrator can set the examination subject, test type questions and the user's permissions.

4.3 Online examination module

Normal users (students) can make exam online. Candidates can login system using account and password, in order to reflect the fairness of the exam, the kinds of questions randomly disrupted secondary reorganization distributed to candidates. This makes the content of the examination had the same to all candidates but with different order of items.

4.4 Manual test modules

Administrator can make basic property settings for test paper according to the different test objects and add paper strategy papers.

4.5 Test management module

Mainly according to the assessment program, set the task type, test quantity, difficulty factor, automatically and quickly making test paper, and perform related operations such as editing on paper. Examination paper consists of exam papers and homework exercises paper.

4.6 Item information management module

To evaluate history papers and analysis the exam, and then to modify the test information for the next exam. Administrators import questions and management database, see question distribution statistics.

4.7 Marking management modules

Mainly making judgment of the papers submitted by the candidates, there are two parts that is to say automatic marking and artificial modification.

4.8 Query management module

Administrators can view, analyze, export and other operations, and can also provide with score search and other functions for students. Students will be free to check your record after participating in the examination.

4.9 News management

Administrators can update, delete and publish news.

5 System implementation

In using Microsoft Visual Studio 2015 and SQL Server 2014 back-end databases, occupational skill testing question bank management system would be implemented for mechanical fitter, machine lathe operator, vehicle maintenance, automobile assembly workers and other categories of professional skills identification papers. Test Bank included alternative questions, choice, fill in the questions and comprehensive question and integrated application 100 questions each, through specific experiments and found that ,within a certain time, advanced particle swarm optimization algorithm can generate optimal solution to automatically generate moderate difficulty, paper covering range of knowledge, student test scores is good using this papers.

Acknowledgments

This work has been supported by Project of the Program of National Natural Science Foundation of China under the Grant Number 11772228 This work was supported by innovative projects in Tianjin University of Technology and Education of china under the Grant Number 201610066023

References

4. Duan Hong. Intelligent test paper composition studies using multi-objective Particle Swarm algorithm[D]. Northeast Normal University 2013


<table>
<thead>
<tr>
<th>Hierarchy of difficulty</th>
<th>Easy</th>
<th>little Easy</th>
<th>Middle</th>
<th>little Difficult</th>
<th>Difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td>0.0-0.2</td>
<td>0.2-0.4</td>
<td>0.4-0.6</td>
<td>0.6-0.8</td>
<td>0.8-1.0</td>
</tr>
<tr>
<td>Degree</td>
<td>0.1</td>
<td>0.3</td>
<td>0.5</td>
<td>0.7</td>
<td>0.9</td>
</tr>
</tbody>
</table>

| Distinction | 0.4 above | 0.3-0.39 | 0.2-0.29 | 0.2 below |
| Evaluation  | good | qualified | Barely changes | Worse, should be eliminated |

Table 1. Hierarchy of Difficulty

Table 2. Degree evaluation criteria