

# The Online Learning and Monitoring System Model of MOOCs

Qi Yang

Tianjin Open University, 1 Yingshui Road, Tianjin 300191, China

yangqi@163.com

**Keywords:** MOOCs, Online Learning, Genetic Algorithm, Monitor

**Abstract.** MOOCs is the rise of a kind of education teaching mode, which has a big difference with the traditional teaching mode. Online learning is also a specific application of modern computer technology in network practice. Genetic Algorithm(GA) is a global optimization algorithm to simulate the life evolution mechanism and achieve random search, which has a very strong calculation ability. In this paper, GA is put forward for MOOCs to monitor the data in the process of online learning.

## Introduction

In recent years, MOOCs is the rise of a kind of education teaching mode, which has a big difference with the traditional teaching mode. Its teaching characteristics of large-scale, online and open breaks the limit of the traditional school which depends on the fixed place of teaching. Under the premise of the condition that has access to the Internet, more than ten thousand people or more people can participate in the study. Online learning is the inevitable trend of network development, which is a form of distance education. Online learning is also a specific application of modern computer technology in network practice. It takes full advantage of the existing network resources. It can transcend the limit of time and space, break the traditional teaching mode and manifest a personalized education through developing an online-learning system. It can not only optimize the resources of network education, popularize education for all people, improve the overall quality, but also can greatly reduce the cost of education, improve the quality of teaching and meet the needs of the learner's personalized features. Therefore, it has a very broad social value.

Genetic Algorithm(GA) is a global optimization algorithm to simulate the life evolution mechanism and achieve random search, which has a very strong calculation ability. Its solving process is simple through high efficient and adaptive control and search for encoded space parameters, to ensure the optimum solution. It is an universally adaptive search method for solving optimization problem, which is one of the key technology for intelligent calculation in the 21st century. In this paper, GA is put forward for MOOCs to monitor the data in the process of online learning.

## Real-Time Monitoring and Control of Learning Process

Real-time monitoring and control mainly records the length of the online learning time of learners, the number of discussion posts, topic selection etc.. According to whether it is real-time between the learning activity and communication, it can be divided into synchronous communication and asynchronous communication. The former is mainly based on longed for sorting. It firstly solves the problems in the top position of MOOCs online, and then to discuss and study. In view of the problem which can't be replied in a timely manner, asynchronous communication mainly uses the manner of EMAIL and BBS to make a delay answer. Different monitoring measures for different communication way, especially to effectively monitor and track for synchronous communication information, it can effectively grasp the learning status and progress of learners.

The framework of MOOCs education online learning and monitoring is shown as follows in figure 1.

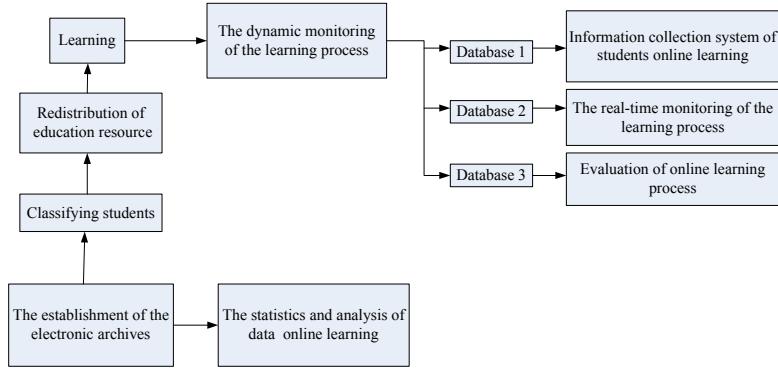


Fig.1: The Framework of MOOCs Education Online Learning and Monitoring

### The definition of Related text

**User Mode.** Learners interest model uses the vector space model(VSM) to express, and then it can be expressed a n-dimensional vector in formula (1).

$$User_i = \{(T_1, S_1), (T_2, S_2) \dots (T_n, S_n)\} \quad (1)$$

Among them,  $T_i$  is the personalized feature information of learners, which is given in the formula (2), S is the relevance of the learners' interest and the feature item.

**Grade Criterion.** Through  $Q$  fuzzy numbers, learners grad for the corresponding learning resources and give the score.  $P_i$  is used to measure the correlation of the text  $X_i$ , and the fuzzy number is converted into accurate number as learners' feedback information for learning resources.

$$E_{i,j} = \frac{P_i + 1}{Q}, 1 \leq i \leq m \quad (2)$$

**Sparseness Calculation of Matrix of Learners-Learning Resources.** The minimum limit value of the sparseness of matrix of the learners-learning resources is predefined to be  $\theta$ , and the actual situation is calculated by the formula 3.

$$Sparseness = \frac{EvaluateNumber}{LearnerNumber \times ResourceNumber} \quad (3)$$

Where, *EvaluateNumber* is the number of the evaluation of learners for learning resource, *LearnerNumber* is the number of learners, *ResourceNumber* is the quantity of resources.

**The Similarity Between the User Model.** In order to use the learning resource selection experience of other users for reference to optimize the model itself, it obtains through calculating the similarity between user model, that's, Is the evaluation function of first goal user model, that's,

$$\Phi_1(X) = Similarity(User_{ik}, User_{jk}) = \frac{\sum_{k=1,..,n} S_{ik} S_{jk}}{\sqrt{\sum_{k=1,..,n} (S_{ik})^2 \sum_{k=1,..,n} (S_{jk})^2}} \quad (4)$$

**Gene Efficiency and Genetic Strain.** Gene efficiency is used to make sure whether the performance play a role in the process of evolution. According to the principle that the user model participate in the evolution is consistent with the information feedback by the user grading, then the gene does not appear in the text and its performance reduces.

Gene performance is different from different user, in order to distinguish them, genetic strain is used to expressed here.

**Chromosome Trust.** If genes of chromosome C is from the User, the User Chromosome-Trust is calculated by the formula (5).

$$\Phi_2(X) = ChromosomeTrust_c = \frac{C_1 * GeneEfficiency_1 + C_2 * GeneEfficiency_2 + \dots + C_{UserCount} * GeneEfficiency_{UserCount}}{C_1 + C_2 + \dots + C_{UserCount}} \quad (5)$$

where  $C_i$  is the number of genes of the user  $i$ ,  $GeneEfficiency_i$  is the gene performance of the user  $i$ .

**Lifecycle Credit.** Lifecycle credit is used to express the number of genetic operators of user model evolution that chromosome participates in. Where the formula of lifecycle credit of the Chromosome is defined as follows.

$$\Phi_3(X) = \begin{cases} LC_c = \frac{ChroOpCount_c}{\max_{i=1, popsize}(ChroOpCount_i)}, & \max(ChroOpCount_i) \neq 0 \\ 0, & \text{else} \end{cases} \quad (6)$$

**The Mathematical Model of User-Model.** According to the definition of multi-objective optimization problems, combined with the reality of the user-model, we describe the user-model by the following formula based on inequality method, that's,

$$\begin{aligned} \text{Minize } & \lambda_f(X, \varepsilon_f), 1 \leq f \leq \pi \quad \pi \text{ is the number of objects} \\ \text{s.t. } & \Phi_i(X) \leq \varepsilon_i, i = 1, 2, 3, 4 \end{aligned} \quad (7)$$

### The Elements of Personalized Learning and the Structure Model

The user information mainly includes the learner's static information and dynamic information. And static information is the basic information of the learners, such as gender, age and archives, etc.. Dynamic information refers to the ability target information and the dynamic collection of the learning resource information in the process of learning.

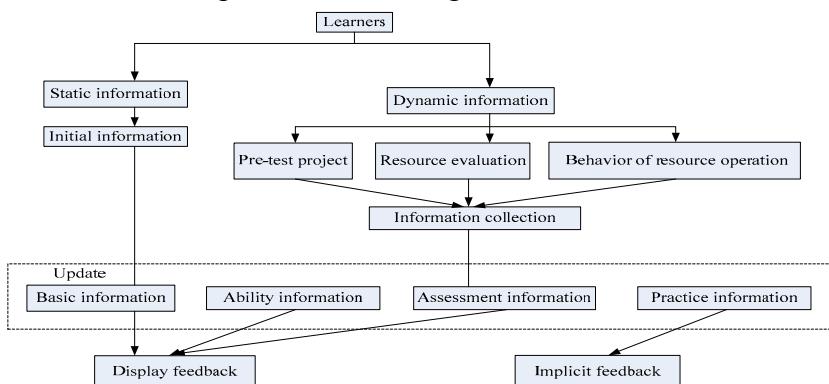


Fig.2: The Construction and Update of Character Main Factor

**Construction of the Initial Study User Model.** Before the evolution of the user model by using genetic algorithm, we must carry on the user's identity to determine and establish the initial user model of learning. Then to judge whether the users use account and password to log in. Through access to the database information of history, we can determine whether the user is a new user or not. For new users, to solve the problem of cold start problem, the method of model based on trust relationship combined with the model based on collaborative filtering of user similarity is used for resources recommendation for the reference users. And for the old user, the steps are divided into the following.

- Step1: Gain the Information of User Evaluation Matrix of Learning Resources;
- Step2: Gain the ability target information through the score of the learners' pre-test project;
- Step3: The Mining of Learners' Implicit Behavior;
- Step4: Express the User Interest Model

**Genetic Operators of User Model in the Evolution.** (i) Initialize the population to randomly generate the matrix whose scale is MXN and then form initial population, m is the current population size, n is the length of the individual;

(ii) Good individual should be able to better reflect the learners' interest preference, therefore, the individual evaluation function should combine with fitness, lifecycle credit and chromosome trust, the value of evaluation function is set to

$$Value(C) = \alpha * Fitness_c + \beta * ChromosomeTrust_c + \gamma * \left( \frac{1}{LC_c} \right) \quad (8)$$

(iii) In the process of genetic encoding, every chromosome represents the teaching resources. Where one gene is on behalf of each feature and its weight, the length of the chromosome may not be the same and the final dimensions of text vector are uncertainty. After overall consideration, we use the dynamic linked list to store individual in order to realize the demand of the length of chromosome to be variable.

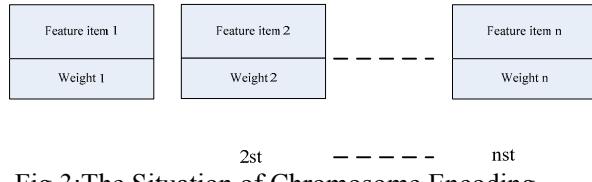
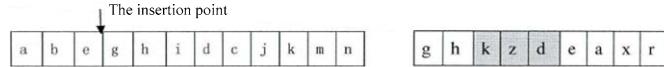


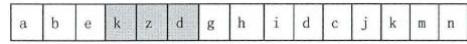
Fig.3:The Situation of Chromosome Encoding

(iv) Due to the set of individual genes composed of feature and its weight, here we adopt the insertion crossover operator. First, we randomly select the insertion point from the father body and the gene fragment. After that the gene fragment inserts the father body, we should delete the duplicated gene. For example,

Step 1: Randomly select male parent, and determine the insertion point and gene fragment.



Step 2: Choose the disjoint genome to insert into the male parent.



Step 3: Remove duplicate genes. Finally the results are shown as follows:



Fig.4:The Schematic Diagram of Crossover Operation

(v) Through a certain variation rate  $p$  and mutation operator to select genes on chromosomes, and the characteristics of the gene is changed. It can keep the population diversity and comfortably improve the searching efficiency of GA.

## Experiment of the Evolution Algorithm Flow of User-Model Based on Genetic Algorithm

First stage: Building the current user's initial interest model. Input data, the text set in a particular field is taken from the MIT-Principle of Data Mining, the page 5 to page 105. The content of each page is used as a text, then the number of this text set is 100. Text is named p00.txt, p01.txt, ..., p100.txt.

The output data: the matrix of text set is expressed as mm.txt.

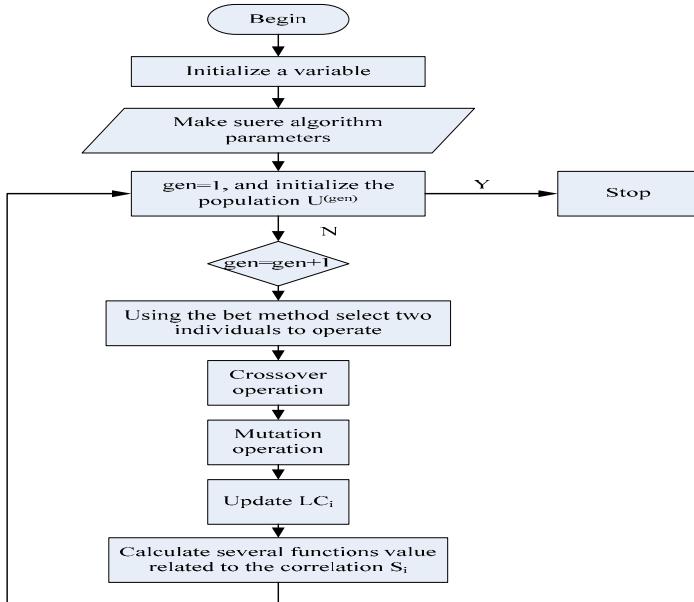


Fig.5:The Frame of User Model Dynamic Updating Algorithm Based on the Genetic Algorithm

The second stage, that's, the stage of the current user-mode. The specific algorithm is as follows.

Step 1: Determine the various parameters of genetic algorithm, that's, the maximum number of iterations g, the downtime threshold e, population size n, crossover probability Pc, the mutation probability Pm. The size of the population is set to 50 chromosomes, other parameters are set as the table 1.

Tab. 1: Parameter Settings

Parameter	Value
crossover probability	0.7
mutation probability	0.2
$\alpha$	0.7
$\beta$	0.2
$\gamma$	0.1
$\theta$	0.1
$\mu$	0.6
$\nu$	0.3
The number of genetic iterations	200
Downtime threshold	0.0001

Step 2: Chromosomes are produced by the way of roulette, cycle, for i= 0 to popsize, initialize the population U(t).

Step 3: Implement the gen=1, where gen is the number of iterations.

Step 4. While gen <g, do the following:

(i)In a population a certain probability P is used to randomly select the individual S and make the crossover operation, and to change the genetic lineage information of the individual.

(ii)From the individual through the crossover operation, using a certain probability P? to randomly select the individual S(gen). And according to the definition of the mutation operator, to make a mutation operation and change an individual's genetic pedigree information.

(iii)If the individual has made the crossover and mutation operation, then update the LCi.

(iv)According to the definition of chromosome trust, to calculate ChromosomeTrusti. Then according to the definition of gene efficiency, to calculate each gene efficiency of each chromosome, among them, Ki<GeneCount. According to the definition of evaluation function, to calculate the value of fitness function of each chromosome.

(v)Implement the selected operation. According to the proportion information of the evaluation

function value of each chromosome, the next generation of chromosome is produced by the way of roulette. Obviously, the Value() is bigger, and the probability of producing the next generation is bigger. In order to accelerate the convergence speed of the algorithm, the individual that the Value() of each generation is bigger can be forcibly inherited to the next generation, and it is not limited to the selection process.

(vi) At this point we get the next generation of population after genetic operation. If the average value of the new population does not change again, then stop; If the number of the genetic iterations is over the set value, then the genetic operation fails.

(vii) gen=gen+1;  
end while

## The result of the training text

To train with the terms weight, data and mining, then the output result is shown in figure 6.

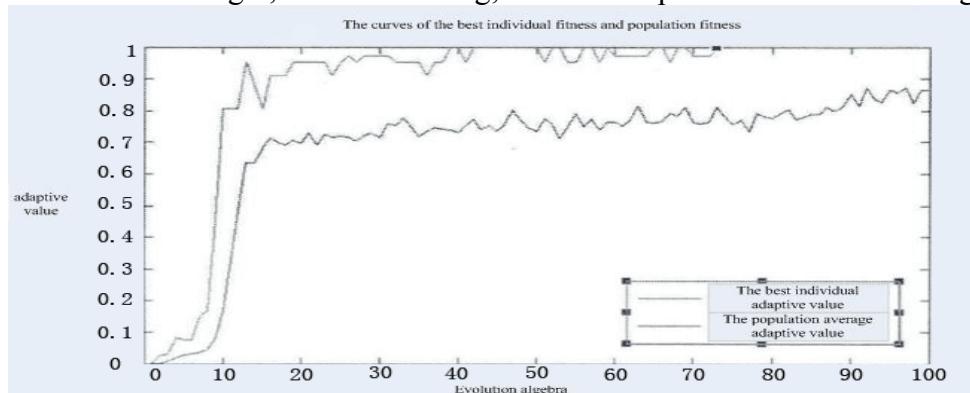


Fig.6:The Best Individual Adaptive Value

## References

- [1] Zheng Jinhua. Multi-objective evolutionary algorithm and its application[M]. Beijing: Science Press, 2007.
- [2] <http://bbs.paidai.com/topic/104810>? 2012. 6.
- [3] Yang Manfu. Research on college students' interactive learning mode based on two kinds of social software[D]. Guangzhou: South China normal university, 2005.
- [4] Laura Pappano. The Year of the MOOC-The New York Times. November 2 2012.
- [5] Irvine V Code J Richards L. Realigning higher education for the 21st-century learner through multi-access learning [J]. MERLOT Journal of Online Learning and Teaching 2013 2.
- [6] Watters, Audrey. (2013). Hack Education, Hack Education Weekly News: MOOCs and Anti-MOOCs. From the Internet: <http://www.hackeducation.com/2013/05/10/hack-education-weekly-news-5-10-2013/>
- [7] Johnson, L., Adams Becker, S., Cummins, M., Freeman, A., Ifenthaler, D., and Vardaxis, N. (2013). Technology Outlook for Australian Tertiary Education 2013-2018: An NMC Horizon Project Regional Analysis. Austin, Texas: The New Media Consortium.