# Abstract

With the rapid development of network technology, many universities have established distance education system to meet the needs of students who cannot easily attend universities. Distance education includes many different aspects, such as the teaching system and examination system, etc. The online examination system is a very important aspect.

This thesis for the first time investigates how recommendation algorithm can be employed to improve examination systems. This system developed in this project is divided two parts, which include real exam and practice exam. In the real exam, the teacher sets an exam and student takes part in the exam. In the practice exam, student can set an exam and take part in the exam. If the student sets an exam for the first time, the system will randomly choose questions for the student. If the student has already set the exam, the system will recommend his/her interested questions. In order to achieve recommendation, I chose slope one algorithm, which has the best accuracy for recommendation. The thesis focuses on the design and implementation of recommendation question and the design of a whole system, including system analysis, database design, recommendation algorithm and system debugging. This system is based on Browse/Server model, using the latest ASP.NET 3.5 framework to achieve. It is simple, easy to use and more secure. Compare with the traditional examination process, the system has reasonable framework design, which includes six modules: authentication, questions bank management, online testing, marking management (computer automated marking and manual marking), practice exam and score management.

Key words: Recommendation algorithm, online exam system, Slope One algorithm, Genetic algorithm

# Table of content

# Chapter 1

# Aims & Objectives

Although distance education has developed vigorously, currently examination method is still traditional in most universities, which can be called 'pen-and-paper method'. The method includes four elements: question papers designed by teachers, question answered by students with pens, marking by hand, and paper analysis. Apparently, the traditional method of examination increases the workload of teachers, and is not suitable for modern distance education, while online examination system can greatly improve work efficiency.

The aim of this project is to develop an Online Examination System based on Recommendation algorithm. Teachers will use this system to evaluate each student objectively. Students will visit this system through intranet or internet. They will follow the instruction from the online system to finish self-test. This system developed in this project is divided two parts, which are real exam and practice exam. In the real exam, teacher sets an exam and student takes part in the exam. In the practice exam, student can set an exam and take part in the exam. If the student is first time to set an exam, the system will randomly choose questions for students. If the student has already set the exam, the system will recommend his/her interested questions. The system can recommend questions to students which are might be interested in.

Objectives:
1) Develop an online examination system
2) Make it secure
3) Automatically generate fair exam paper
4) Practice module with recommendation system
5) Test the system

# Chapter 2

# Background & Related technologies

## 2.1 Software models of internet applications

There are two main approaches to development of internet applications, which are C/S (Client/Server) model and B/S (Browser/Server) model. C/S model needs clients to install special client software and operating environment. First, the workload of installation is very much. Followed by any computer problems, such as viruses, hardware damage, these are required to install or maintain. In addition, as system software upgrades, each client needs to re-install. The costs of maintenance and upgrade are very high. The advantage of B/S model is that clients do not need to maintain, because all clients are browsers. No matter how big the size of the user, it will not add any workloads of the maintenance and upgrades. All of the operations are only for the server. If it is off-site, we can remote maintenance and upgrades. Considering safety and maintenance, B/S model is better than C/S model for the proposed application.

## 2.2  Existing online examination

The United States took the lead using the computer technology in education and now has entered a mature stage. Web-based platform is used in many famous international

computer certificate tests and foreign language tests, such as Microsoft's certification exam, Sun Java certification exam, Cisco Certified Network Associate (CCNA) and so on. Candidates can apply for the test in the authorized test centers, after determining the test of time, candidates can take the test. [14]

Remote test service has also become a quite promising industry. For example, *Sylvan Prometic* in the United States is a professional company for examination service. The company sponsors Computer Aided Testing (CAT). Through the network, exam questions can be sent to every examination room around the world. The type of exam questions provided by the company is multiple choices. Candidates only need to log on to its website in advance and then they can take the exam. They will know the results of test immediately after the examination. At the same time, the results will be immediately sent to the corresponding institutions. Most of certification institutions (Microsoft, Cisco, Oracle, etc.) entrust the company to carry out online exam.

In China, distance education also contributes to a number of online exam systems, such as remote test system developed by Shanghai Communications University Education Centre, common examination system developed by Beijing Jeffrey Co., Ltd, PowerExam common examination system developed by Beijing Tianshi Co., Ltd, Neoexam online examination system developed by Chengdu Science and Technology Development Co., Ltd. and so on.[15]

Most of online examination systems currently use C/S (client / server) model. The model requires installing the exam software in both the client side and server side. Client side software includes examinations and auto-mark module, responsible for displaying the test paper and communicating with the server. Server side software includes paper management and test management module, responsible for maintaining the question bank, auto-generating paper, collecting the results of exam and statistics the results etc. For example, National Computer Rank Examination (NCRE) system in China uses C/S model. The model has a great dependence on the client software. It is not convenient for system maintenance, and maintenance cost is high. In a word, it is not suitable for a large test.

In recent years, with the development of ASP.NET, JSP and other web interactive technologies, online examination system based on B/S(browser/server) have become more popular. This model uses a database and web technologies. The browser should be installed in the system in the client side, but there is no need to install any additional software. It can generate questions, mark the paper and collect results automatically, while the system is very easy to be maintained and upgraded. Candidates just need to login to the system and take the exam through the Internet, no longer restricted by geographical area and time. At present, B/S model is used in certification exams such as Microsoft's various certification exams. This approach meets the needs of modern education, particularly of current mature distance education. It represents the future direction of development of examination systems. Modern

online education mainly uses Web-based commercial course management software such as WebCT (e.g. used by University of central Lancashire) and Blackboard (e.g. used by University of Bristol).

The organization of the online examination will be divided into three types:
1) Full open: this type is used for self-assessment of student. It is open, so everyone can access to the system. Normally, the teacher does not need to monitor the exam, and the test can be taken at any time and any place. Students can get results of test immediately after the test. Generally, it is free.
2) Semi-open: this type is used for certification exam, such as CCNA and NCRE that I mention before. Candidates should apply for exam and are required to take the test in designated examination room. In semi-open mode, arrangement of the exam is relatively free. Candidates can select any exam in required time-range. Normally, there are a number of teachers responsible for monitoring the exam; candidates need to pay fees for exams. For example, CCNA examination fee is 120 pounds.
3) Closed type: this type is used for exams in enterprises or universities. Candidates must take the exam in required time and place, and there are a number of teachers responsible for monitoring the exam.

The above three types of online exams only use browsers, and exam software no longer needed. All the information and test results are saved on the server.

## 2.3    Microsoft SQL Server database technology

Microsoft SQL Server 2005 is a relational database management system with high performance. As a member of Windows database family, SQL Server relational database management system can meet all types of enterprise customers' requirements. [11] According to customer's demand, SQL Server 2005 has significantly improved in scalability, usability and manageability. [11]

SQL Server 2005 uses client/server (C/S) model. All workloads are decomposed into the server and client. The client application is responsible for Business Logic Layer and provides data to users. [11] Generally, it run in one or more client machines and also can run on the server. Server is responsible for the management and allocation of the available resources, such as memory, network bandwidth and disk I/O (input/output) etc.

SQL Server 2005 is a part of Relation Database Management System (RDBMS). It is responsible for managing the database structure, which mainly include the maintenance of the relationship between the data in the database so as to ensure the data accuracy and recovery of all data once the system fails.

# 2.3.1 Features and functions of Microsoft SQL Server 2005

SQL Server 2005 extends the performance, reliability and accessibility based-on previous version (SQL Server 7.0). [12]

SQL Server 2005 has the following new features: [12]

1) Support Extensible Markup Language (XML). SQL Server 2005 allows the use of XML to configure the system and exchange data. Developer can use XML to insert, update and delete data.

2) Safely access the database by web browser. SQL Server 2005 provides a browser through the firewall security to access the data.

3) Scalability. In order to improve the scalability and manageability of a large table, SQL Server 2005 allows using partitions in data warehouse. [13]The partition will help to reduce query time, improve load time and maintainability of the database. As the tables become very large, partitioning can divide tables into smaller. Thereby, it improves maintainability of the database. If we have multiple CPU in PC with a large table, the partition of the table can achieve better performance by parallel operations.

4) Maximum uptime and reliability. It allows keeping the database online status in database maintenance. SQL Server 2005 enables fast backup, and sends the log file to the database maintenance plan.

# 2.4 Genetic algorithms for setting fair test paper

## 2.4.1 Principles of setting test paper

A piece of test paper is not a random combination of questions. Many people think that the process of paper-making includes: to collect a large number of questions, to put them into question bank, and to classify them, then randomly select questions to compose a piece of paper. However, such process cannot generate exam paper with balanced difficulty [16]. The real purpose of developing online examination system is to generate objective and scientific papers [17]. To achieve the goal, the following four principles should be stressed.

1) Principle for goals: Each course should include teaching objectives. The

examination is a method for achieving teaching objectives. It can measure learning effect of students and test whether teaching objectives are achieved.

2) Principle for all sidedness: the paper should include all knowledge points. When it is established, the question bank should cover all knowledge points. Each of knowledge points should be covered by the corresponding test questions, and there should be different types of questions in the same point.

3) Principle for significance: The test paper should highlight the most important part of the course and the essence of knowledge.

4) Quantitative principle: In the process of generating test paper, questions should be strictly allocated in accordance with the teaching requirements. In addition, the difficulty of test questions should be suited to the level of candidates.

## 2.4.2 Constraints of setting test paper

Questions should be designed according to inherent characteristic parameters, which are called constraint. By establishing the constraint, the system can generate exam paper module. There are seven following type of constraints [17]:

1) Types of questions: multiple choice, true-false, fill in blank, short answer, etc.
2) Respective chapters: Questions should be corresponding to the chapters of the course in the questions bank.
3) Knowledge points: The specific exam questions should include the knowledge points.
4) Difficulty coefficient: Difficulty coefficient is calculated by following method. Let the difficulty of $j^{th}$ question be $p_j$:
   For subjective question, the difficulty coefficient is:

$$p_j = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}x_{ij}}{w_j} = 1 - \frac{\overline{x}_j}{w_j}$$

The $N$ denotes the total of candidates. The $x_{ij}$ denotes score of $i^{th}$ student at

the $j^{th}$ question. The $\overline{x}_j$ denotes the average score of $j^{th}$ question. The $w_j$

denotes the values of $j^{th}$ question.

For Objective question, the difficulty coefficient is:

$$p_j = 1 - \frac{n_j}{N}$$

The $N$ denotes the total of candidates. The $n_j$ denotes the number of student who correct answers the $j^{th}$ question.

The value range of $p_j$: [0-0.4] hard, [0.4-0.7] medium, [0.7-1] easy.

5) Teaching requirement: It includes memory, understanding, simple application and integrated application.

6) Score: the test score of the student is determined by the number of correct answers, exam time and item difficulty.

7) Utilization ratio: it refers how many times a question has been selected by the system. Out of considering fairness, the system should control the utilization ratio during the generating exam paper.

## 2.4.3   Automatic generation test papers

Automatic generation of test paper means to select questions from the Question Bank and to compose a paper according to user conditions and algorithm. Thereby, the selection of algorithm is important, which determines whether maximized fairness is satisfied for every student.

To achieve the automatic generation of test paper, there are three methods:

a. Randomizing. [6] Randomly picking questions with some parameters is a usual method to make a test paper. This process repeats until the test paper generation is completed or one cannot select questions from the Question Bank to meet the constraints.

b. Backtracking.[7] "Backtracking is a general algorithm for finding all (or some) solutions to some computational problem, that incrementally builds candidates to the solutions, and abandons each partial candidate $c$("backtracks") as soon as it determines that $c$ cannot possibly be completed to a valid solution." (Wikipedia Backtracking) Backtracking uses the Depth-First-Search algorithm. It is suitable for small Question Bank in online examination system.

c. Genetic algorithm.[8] Genetic algorithm is proposed by Professor J. Holland in 1975. It is a computation module which simulates Darwinian biological evolution. It is a method of searching for optimal solution by simulation natural evolutionary process. It is suitable for generating test paper in online examination system. This system used genetic algorithm to generate fair test paper. It is satisfied with intelligent, high quality and high speed to generate test paper. Finally, it is suitable for large-scale online exams.

## 2.4.4   Summary of genetic algorithm

Biological evolution is an optimized process. It can produce prefect species by environmental selection and holmatogenesis. Genetic algorithm is also an optimized algorithm which is inspired by the idea of biological evolution. The concept of genetic algorithm was firstly proposed by Bagley J.D in 1967. Professor J.H. Holland in the University of Michigan made a pioneering research on the theory and method of the genetic algorithm in 1975. In recent years, genetic algorithm has been widely used in pattern recognition, neural networks, image processing, machine learning, industrial optimal control, adaptive control, biological sciences, social sciences and other aspects. [8]

The basic idea of Genetic algorithm is based on Darwin's *theory of evolution* and Mendel's *theory of heredity.*

The important principle of Darwin's *theory of evolution* is survival of the fittest. He thought that every species would be more and more suitable for living environment. The basic features of each species are inherited to its offspring, but some changes would take place in gene of the offspring, which is called mutation in genetics.

The important principle of Mendel's *theory of heredity is* gene genetic. He thought that heredity exists in cells in the form of the genetic code, and in the chromosome in the form of genetic gene. Each gene has a special place and controls a particular character. Gene mutation and gene hybridization can make offspring to adapt to environment better. After natural elimination, highly adaptive gene can survive.

According to *theory of evolution* and *theory of heredity,* the following concepts will be used: [8] [18] [19]
1) String: In genetic algorithm, there is a binary string and it is correspond with chromosome in genetics.
2) Population: a set of individuals of Chromosome with characteristics are called the population.
3) Gene: gene is an element of string and it represents the characteristics of individual.
4) Gene Position: It means a place where gene is located. Gene position is calculated from left to right by string.
5) Evolution: In the process of biological survival, the species gradually adapt to environment, so biological quality is constantly improved. This process is called evolution.
6) Fitness: fitness represents adaptive level of an individual. In the study of the genetic and evolutionary biology, the biologists use the term (fitness) to measure the adaptive level of a species in the environment. High fitness of the species will get more chance to breed, but low fitness of species will gradually die out.

# 2.4.5   Structure of genetic algorithm

The basic structure of the genetic algorithm is shown in Figure 1.
1) Set generation $k=0$ and produce initial population
2) Calculate the fitness of each individual in initial population
3) Randomly select two individuals from initial population, and make a mutation operation with two individuals, then get a new generation, let $k=k+1$
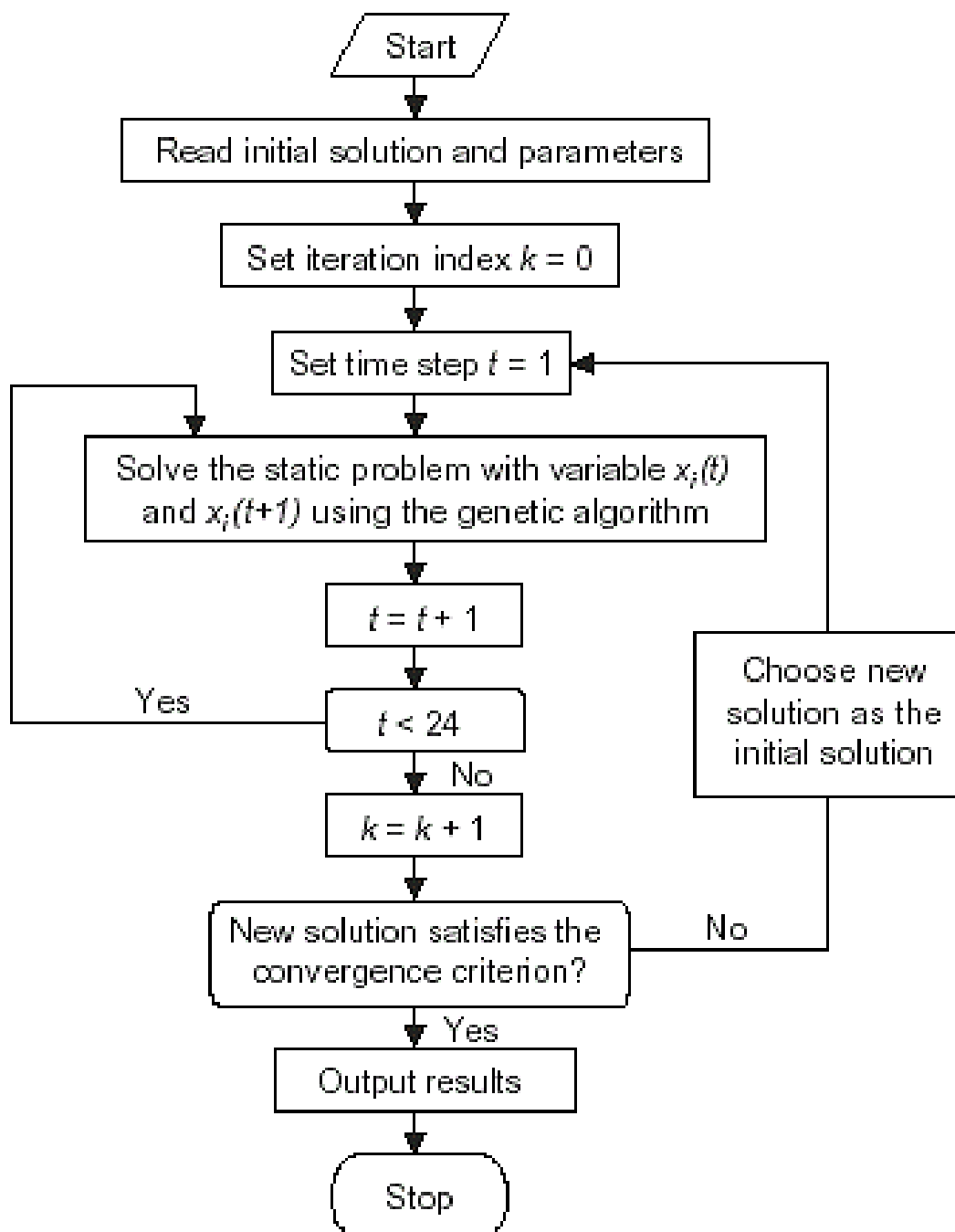4) If the terminal condition is satisfied, then the algorithm is terminated; otherwise, go back (2).

Figure 1 Structure of genetic algorithm [18]
Source: http://www.atsdr.cdc.gov/hs/dover2/doverhtml/appendix/appe.html

## 2.5 Recommendation system

Recommendation system uses statistics, artificial intelligence and data mining techniques to analyze behavior of visitors. It can generate recommendations to help users access product information in which they are interested. It can guide the user's buying behavior and create substantial profits for users. To recommend exam questions for each student, I used this technique in practice part of my online exam system.

### 2.5.1 Main recommendation algorithms

The recommendation algorithm is the core and the most critical part in the recommendation system. It largely determines the performance of recommendation system. Currently, the main recommendation methods include: content-based recommendation, collaborative filtering, association rule-based recommendation, utility-based recommendation, knowledge-based recommendation and hybrid recommendation. The collaborative filtering algorithm is particularly widely used and there are user-based collaborative filtering and item-based collaborative filtering. As for user-based or item-based collaborative filtering, there are two data acquisition method which are implicit method and explicit method. Explicit method needs user input some information, such as rating the product. Implicit method does not require users to input evaluation data, but the system makes evaluation for users according to the behavioral characteristics of the users.

### 2.5.2 Content-based Recommendation

Content-based recommendation is a continuation of the information retrieval technology [4]. "Information retrieval is the science of searching for document, as well as that of searching relational database and the World Wide Web." (Wikipedia information retrieval) Content-based recommendation method recommended products based-on the content information without the user's evaluation of the items. It needs to use machine learning approach to get user's interest information from the content. For example, I have chosen a book *Data Mining, Second Edition: Practical Machine Learning Tools and Techniques, Second Edition (Kindle Edition)* in Amazon.com. The Amazon.com system recommended some books are following:

Figure4 Amazon.com system recommend books (source:www.amazon.com )

In content-based recommendation system, the item or object is defined through the relevant case feature. The system calculates the match-level between users' data and the predicted items for current user. The user data model depends on the learning method, commonly used with decision trees, neural networks and vector-based methods. Content-based users' information needs the users' historical data. User data model may change with the users' preferences.

The advantages of content-based recommendation method are as follows [5]:

1) It does not require other users of the data. No cold-start problems and data sparsity problems. "Cold-start is a potential problem in computer-based information systems which involve a degree of automated data modeling. Specifically, it concerns the issue that the system cannot draw an inferences for users or items about which it has not yet gathered sufficient information."(From Wikipedia).   ""
2) It can recommend for users with special interests.
3) It can recommend new items.
4) It can recommend the project by listing the contents of characteristics, which can explain why those items are recommended.

The disadvantages of content-based recommendation method are following:
  1) It requires that the content should be easily extracted into a meaningful feature and that the characteristics of content should have a good structure.
  2) It cannot explicitly get information about other users' judgment.


## 2.5.3   Collaborative Filtering Recommendation

"Collaborative filtering (CF) is the process of filtering for information or patterns using techniques involving collaboration among multiple agents, viewpoints, data sources, etc." (From Wikipedia collaborative filtering) It generally uses the nearest neighbor technique, by using the user's historical information, to compute the similarity between the users. Then system uses the neighbor users' evaluation of

products (examination questions) to predict the target user's preference. According to this preference, systems can recommend products (e.g. examination questions) for users [1][5]. An advantage of the collaborative filtering is no special requirements for recommended object. It can deal with unstructured complex object, such as music and movies.

Collaborative filtering is based on the assumption that: A user can find the interested content of the product through finding similar interests among other users. Then system recommends these products (examination questions) for the users. This idea is very easy to be understood. In daily life, we often carry out some options according to a good friend's recommendation. Collaborative filtering exactly used this idea in recommendation system, so this is a process of multi-user "collaborative" to filter items.

Collaborative filtering recommendation system recommends automatically products from the perspective of the user. That is, a user acquires recommendations from buying mode of other users. Users do not need to search for recommendation information to suit their interests, such as filling out survey forms.

Compared to content-based filtering method, collaborative filtering recommendation system has the following advantages [5]:
1) It can filter some information which is difficult to be filtered by automatic content analysis of machine, such as art and music.
2) It can share the experience of other users. It avoids the incomplete or inaccurate content analysis. Besides, it can filter complex or elusive concept, such as information quality and personal taste.
3) It can recommend new information. Compared to the content-based filtering recommendation system, the content-based filtering recommendation system recommends products which are familiar with the contents of the user, but collaborative filtering recommendation system can find potential interests of the users.
4) It can effectively use other similar user's feedback and accelerate individual learning.

Although collaborative filtering recommendation system is significant as a typical application, there are still many problems to be solved. The most typical problems are sparsity and scalability.

## 2.5.4 User-based Collaborative Filtering

Collaborative filtering is also classified user-based collaborative filtering and item-based collaborative filtering. To build a user-based collaborative filtering system typically requires three steps.

Step one. We need to collect information that can represent the users' interest. The traditional way commonly uses scoring system. The most famous example is MovieLens.org.



**Customer reviews & ratings (22)**

| Product summary | | | |
|---|---|---|---|
| Overall Rating | ★★★★★ | 5 Stars | Excellent ⓘ |
| Picture Quality | ★★★★★ | 4.9 Stars | Good |
| Sound Quality | ★★★★★ | 4.7 Stars | Good |
| Connectivity | ★★★★★ | 4.9 Stars | Good |
| Design | ★★★★★ | 4.8 Stars | Good |

Figure 5 customer reviews & ratings (source:www.argos.co.uk)
Some website often appears "reviews and ratings". (Figure 5)

This approach is called "explicit score" method. It has an obvious disadvantage that data collection is difficult. Users often do not want to waste time to make the task for the website. In the actual business systems, even using this method, it will be packaged as a more user-friendly style. There is a more effective method that is "implicit score" method. This method does not require users to input evaluation data, but the system makes evaluation for users according to the behavioral characteristics of the users.

Step two. We need to search nearest neighbor. The starting point of collaborative filtering is finding your interest in the same group of users. The terminology is called "nearest neighbor". The core of nearest neighbor search is computing the similarity of two users. For example, there are user A and user B. The system needs to obtain the score of all items of user A and user B. Then system selects an appropriate algorithm to calculate similarity. Based on score data, system calculates the similarity value of user A and user B. Currently, there are some algorithms including Person Correlation Coefficient, Cosine-based Similarity and Adjusted Cosine Similarity. The details of these algorithms will be discussed in next section. From figure 6, they used data from MovieLens recommender system. There are over 43000 users who have expressed opinions on 3500+ different movies [1]. The following is formula of mean absolute error:

$$MAE = \frac{\sum\limits_{a \in U_{test}} \left| r_{a,j} - P_{a,j} \right|}{U_{test}}$$

The *r* denotes real rating by user. The *P* denotes prediction rating by the algorithm calculated. The *U* denotes users.

The accuracy of adjusted cosine similarity algorithm is better. Usually, we will be based on different data sets to select a different algorithm.

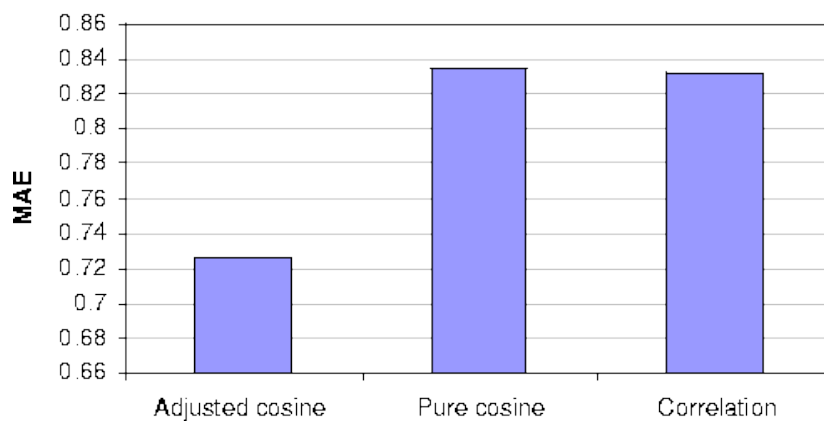**Relative performance of different similarity measures**



Figure 6 Relative performances of different similarity measures, Figure from: ref [1]

Step three. We need to generate a recommendation result. With the nearest neighbor set, we can predict the interests of target user and generate the results of recommendations. Usually, according to the purpose of the recommendations, there are different kinds of recommendation. The most commonly recommendation method is Top-N recommendation and associated recommendation.

Here's Top-N list is different from the "hottest" list. Hot list is generated based on the full data set. It is the same for each person. Top-N recommendation is generated for a single user. It is not the same for each person. Through users' nearest neighbor statistics, select the highest frequency items and these items are not in users' rating as a recommendation results.

As the number of user increases, the amount of calculation of user-based collaborative filtering algorithm is increased linearly. Its performance will be worse and worse. For Web applications, one of the most important factors that affect the user is the speed of response, which hinders the user-based collaborative filtering technology to be used in the actual system. In general, Amazon.com uses Item-based collaborative filtering technology. Along with Amazon's success, Item-based method is also much more

popular.

# 2.5.4.1    Pearson    Correlation    Coefficient    with recommendation system

Pearson correlation coefficient is used to measure the similarity between the two data sets. [1] For example, in a film information website, many users are likely to rate one of the film. Pearson correlation coefficients can be used to help the user to find similar interests, thus the related recommendation is made. The basic idea of this recommendation is that if A, B share similar interest, thus B will probably like the favorite movie of A. Therefore, we can recommend favorite movie of A to B.
Pearson correlation coefficient formula is as follows:

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{(\sum X^2 - \frac{\sum X^2}{N})(\sum Y^2 - \frac{\sum Y^2}{N})}}$$

*X, Y* means a person, which is vector. *N* is amount of items.

Suppose that the film library has three films, both audience A and audience B make score on these films (6 is the highest). A's score was [1, 2, 3]; B's score was [2, 5, 6]. Then the Pearson correlation coefficient of A and B was 0.9608, indicating that the interest of the two persons is very similar.

```
ΣXY = (1)(2) + (2)(5) + (3)(6) = 30
ΣX  = 1 + 2 + 3 = 6
ΣX² = 1² + 2² + 3² = 14
ΣY =  2 + 5 + 6 = 13
ΣY²= 2² + 5² + 6² = 65
N=3
ΣXY−ΣXΣY/N = 30 − (6)(13)/3 = 4
ΣX² − (ΣX)²/N = 14 − 6²/3 = 2
ΣY²− (ΣY)²/N = 65 −13²/3 = 8.667
r = 4 /√(2)(8.6667) = 4/4.16333
   = .9608
```

## 2.5.4.2    Cosine-based Similarity

User rating can be seen as the n-dimensional space. If the user has not rated the item,

the item's score will be set to 0.The similarity between users can be calculated as the cosine of the angle between these two vectors. Let item's rating of user $X$ and user $Y$ in the n-dimensional space be expressed as a vector $\vec{X}$, $\vec{Y}$. Thus the similarity (*sim (X, Y)*) of user $X$ and user $Y$ is:[1]

$$sim(X,Y) = \cos(\vec{X}, \vec{Y}) = \frac{\vec{X} \bullet \vec{Y}}{\|\vec{X}\| * \|\vec{Y}\|}$$

The "$\bullet$"means the dot-product of the two vectors." [1]

## 2.5.4.3    Adjusted Cosine Similarity

In the cosine-based similarity, the metric issue of different user scoring does not be taken into account. By subtracting the average of user's score, adjusted cosine similarity improves these shortcomings [1]. Let $I_{ij}$ be collaborative items' rating of user i and user j. $I_i$ and $I_j$ are items' rating of user i and user j. Thus the similarity (sim (i, j)) between user i and user j is:

$$sim(i,j) = \frac{\sum_{c \in I_{ij}} (R_{i,c} - \overline{R_i})(R_{j,c} - \overline{R_j})}{\sqrt{\sum_{c \in I_i} (R_{i,c} - \overline{R_i})^2} \sqrt{\sum_{c \in I_j} (R_{j,c} - \overline{R_j})^2}}$$

$R_{i,c}$ denotes user $i$ rating on the item $c$. $\overline{R_i}$ denotes the average rating of user $i$. $\overline{R_j}$ denotes the average rating of user j.

## 2.5.4.4    Generate recommendation

Through the above similarity measure, we can get the nearest neighbor of target user. The next step will generate the appropriate recommendation. This method can predict the items' score which user has not rated, and then select the top results of the predicted scores as a recommendation back to the current user. Let *NNu* be the nearest neighbor set of user $u$. Let $P_{u,i}$ be the user u rating for the item $i$. We can calculate $P_{u,i}$ as follows:

$$P_{u,i} = \overline{R_u} + \frac{\sum_{n \in NN_u} sim(u,n) * (R_{n,i} - \overline{R_n})}{\sum_{n \in NN_u} (|sim(u,n)|)}$$

*sim (u, n)* denotes the similarity of user u between user n. $R_{n,i}$ means user n rating for item $i$. $\overline{R_u}$ denotes the average rating of user u. $\overline{R_n}$ denotes the average rating of user

n.

## 2.5.5   Item-based collaborative filtering

Item-based collaborative filtering is popularized by Amazon.com (users who buy x also buy y). Under the context of rating-based collaborative filtering, it was firstly proposed by Vucetic and Obradovic in 2000 [2], and it proceeds in an item-centric manner:
  1) Build an item-item matrix determining relationships between pairs of items
  2) Using the matrix, and the data on the current user, infer his taste

Item-based collaborative filtering is based on the assumption: If most users in some items the scores were similar, the current user's score on these items are also quite similar. Item-based collaborative filtering system uses statistical techniques to find the nearest neighbor of target item. [1] Thereby, we can predict the target item score of the current user according to the score of nearest neighbor. This assumption is consistent with our daily life. Basically, people like bread and they also like milk.
Just like the user-based method, Item-based method requires three steps:
  1) System needs to collect rating data of users.
  2) System needs to search for items of nearest neighbor and calculate the similarity of the items. (The user-based method calculates the similarity of the users)
  3) System needs to generate result of recommendation.

Compared with the user-based methods, Item-based method has improved the scalability of collaborative filtering. In user-based method, with the increase of number of users, the amount of calculation increases linearly. Item-based method calculates the similarity between items instead of the similarity between users. For the item concerned, the similarities between items are more stable. It can complete off-line calculation of the similarity, thus it greatly reduces the online calculation and improves efficiency of recommendation.

For both of content-based and item-based collaborative filtering, there is no verdict for which one is optimal. Since different algorithms to deal with different data were very different. However, there is a great difference using these two algorithms in the concrete realization. That is produced two different approaches, which including online calculations and offline calculations. That is, User-based collaborative filtering algorithm generated recommendations which are online computation, and item-based collaborative filtering algorithm can pre-process a lot of data before generating recommendations, then it can reduce the system pressure.

For both content-based and item-based collaborative filtering, there is no clear conclusion that which one is better, since algorithms to deal with different data are

very different. There are two different approaches, including online calculations and offline calculations. User-based collaborative filtering algorithm generates recommendations by online computation, while item-based collaborative filtering algorithm can process a lot of data before generating recommendations (offline), therefore it can reduce the load of the system.

## 2.5.6   Slope one algorithm

Slope one algorithm is proposed by Daniel Lemire in 2005. It is belong to family of item-based collaborative filtering. It aimed to achieve five goals [21], which including:

1) Easy to implement and maintain: The engineer can easily interpret all aggregated data and the algorithms should be easy to implement.
2) Updated on the run-time: Add a new rating items, it should be an immediate impacted on the predicted results.
3) Efficient query response: query should be faster but it may need to pay the cost of space.
4) Expect little from first visitors: For a score item of a few users, they should get a valid recommendation.
5) Reasonable accuracy: the recommended items should be accuracy within reason.

"Slope one algorithms work on the intuitive principle of a 'popularity differential' between items for users. In order to measure this differential, the simple way is to minus the average rating of the two items. In turn, this difference can be used to predict another user's rating of one of those items, given their rating of the other. " [21] For example: look at figure 2.5.6.
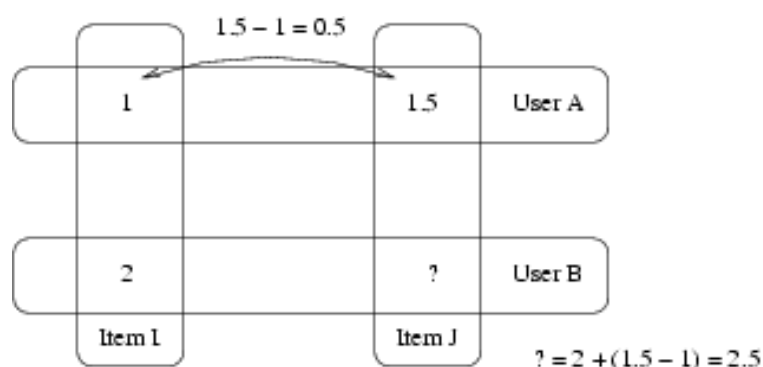


Figure 2.5.6 Example of slope one (source: [21])

User A gave a rating 1 for item I and 1.5 for item J.
User B gave a rating 2 for item I, so how many rating does User B gave for item J?

User A likes item J more than item I by 0.5. We assume user B will also like item J more than item I by 0.5. So we predict the rating of item J is 2.5. Therefore, the answer is 2.5. (2 + (1.5 - 1) / 1 = 2.5).

Here we can see that an advantage of the Slope One algorithm is that user can get a relatively accurate recommendation in a few data, which can be solved Cold Start problem.

However, for realistic application, we need using weight slope one. Since different weight will predict different values. Consider following example:

| Customer | Item 1 | Item 2 | Item 3 |
| --- | --- | --- | --- |
| John | 5 | 3 | 2 |
| Mark | 3 | 4 | Didn't rate it |
| Lucy | Didn't rate it | 2 | 5 |

First, we need to calculate the average difference of item 1 to item 2, which was ( (5-3) + (3-4) ) / 2 = 0.5. The average difference of item 1 to item 3 was 5-2 = 3. Then we predicted the rating of Lucy for item 1 using her rating for item 2. We got 2 + 0.5 = 2.5. In a similar way, we predicted the rating of Lucy for item 1 using her rating for item 3. We got 5 + 3 = 8. So we got two predictions rating of Lucy for item 1. We can't just choose one of them. Hence, we should use weight slope one. "If a user rated several items, the predictions are simply combined using a weighted average where a good choice for the weight is the number of users having rated both items." [22] In the example, we would predict the following rating for Lucy on item 1: ( 2 * 2.5 + 1 * 8 ) / (2+1) = 4.33.

Hence, the formula for weight slope one is:

$$P^{wS1}(u)_j = \frac{\sum_{i \in S(u)-\{j\}} (\mathrm{dev}_{j,i} + u_i) c_{j,i}}{\sum_{i \in S(u)-\{j\}} c_{j,i}}$$

where $c_{j,i} = card(S_{j,i}(\chi))$.

Notation [21]:

i,j means item i and item j. $u_i$ is the rating of this user gives to item i. The S(u) means the subset of the set of items consisting of all those items which are rated in u. The number of items in a set S is card(S). The set $S_i(X)$ is the set of all ratings from a user u $\in$ X such that they contain item i (i $\in$ S(u)). dev means the average deviation of item I with respect to item j. so the formula of dev is:

$$\mathrm{dev}_{j,i} = \sum_{u \in S_{j,i}(\chi)} \frac{u_j - u_i}{card(S_{j,i}(\chi))}.$$

## 2.5.7 Comparison of recommendation algorithm

The following experiment data is from [23]. They have compared to the accuracy of three recommendation algorithms based on MAE (Mean Absolute Error), which are user-based collaborative filtering algorithm, item-based collaborative filtering algorithm and slope one algorithm.

| Number of items | Number of users | Number of ratings | Sparsity / % |
|---|---|---|---|
| 498 | 1177 | 60546 | 10.330 |
| 998 | 1226 | 102482 | 8.376 |
| 1497 | 1198 | 127483 | 7.108 |
| 1994 | 1190 | 146598 | 6.178 |
| 2486 | 1151 | 147732 | 5.163 |
| 2960 | 1198 | 155789 | 4.393 |
| 3341 | 1149 | 151285 | 3.941 |
| 3394 | 1199 | 152451 | 3.746 |

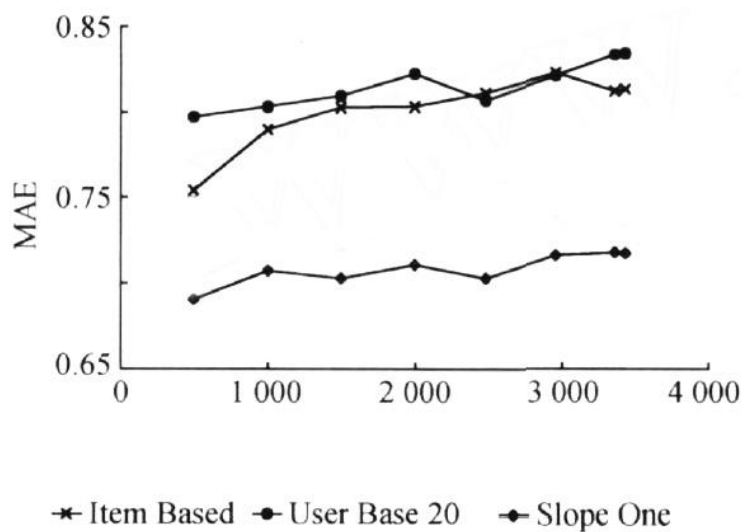Table 2.4.7-1 MovieLens, dataset selected by percentage of items [23]



Figure 2.4.7-1 Comparison algorithm by Items [23]

Using dataset form the table 2.4.7-1, we can see, with the number of items increases, each algorithm's MAE values become larger, actual prediction accuracy will be worse. Slope One algorithm is the best prediction accuracy.

| Number of users | Number of items | Number of ratings | Sparsity / % |
|---|---|---|---|
| 189 | 2841 | 23314 | 4.342 |
| 408 | 3192 | 52896 | 4.062 |
| 602 | 3268 | 78739 | 4.002 |
| 764 | 3295 | 93889 | 3.730 |

| 917 | 3391 | 128146 | 4.121 |
|-----|------|--------|-------|
| 1199 | 3394 | 152451 | 3.746 |

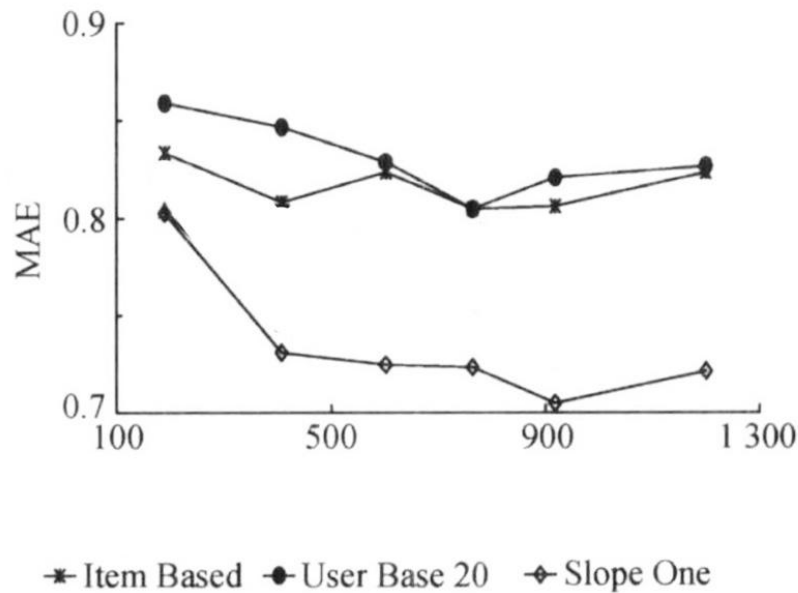Table 2.4.7-2 MovieLens, dataset selected by percentage of users [23]



Figure 2.4.7-2 Comparison algorithm by Users [23]

Using dataset from table 2.4.7-2, we can see, as the number of users increases, besides Item based algorithm, the prediction will be much accurate. Slope One algorithm still is the best prediction accuracy.

# Chapter 3

# System specification & design

## 3.1 Functionality of the system

```
                    ┌─────────────────────┐
                    │ Online Exam System  │
                    └─────────────────────┘
```
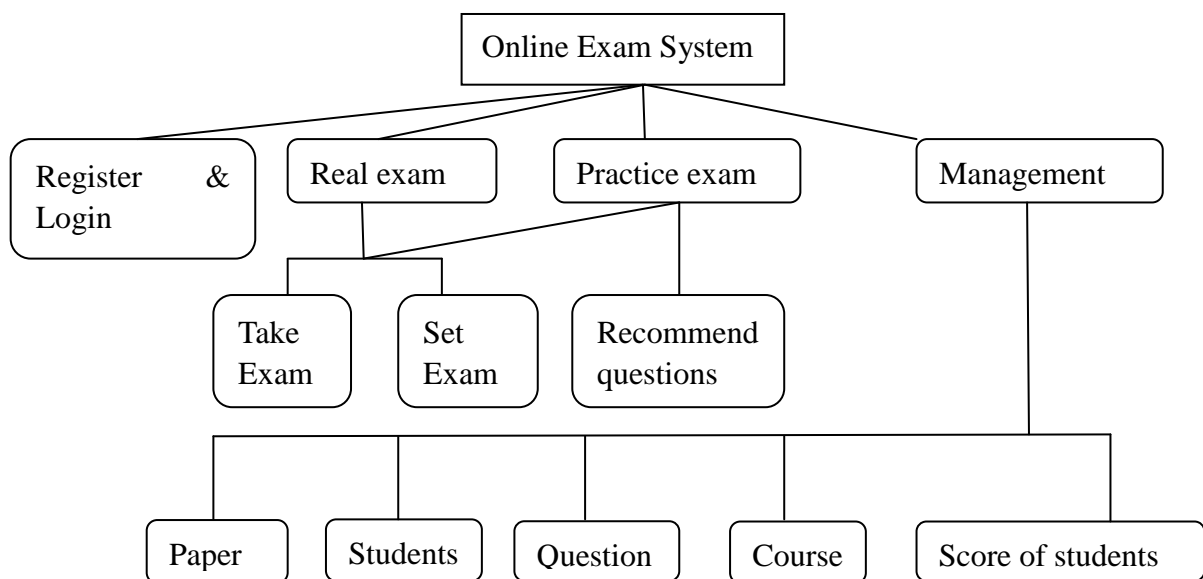
Figure 3.1 functionality of the system

The system is divided into four major parts, which are login module, practice exams

module, online examination module and management module. In the login module, a user who has an account number can directly login the system. Alternatively, a user can register an account number then login the system.

As the special requirements of online examination system, making data security and confidentiality is very important. Confidentiality refers to the user all the information online should be security. Security refers to the user to participate in an event should be a safe process. For example, encrypt the exam paper. Although hacker can intercept the test paper, they haven't got the key to decrypt it. In order to ensure its security and confidentiality, we should ensure that candidates are not cheating. Second, we must strictly control the examinations time.

In the practice exams module, all users (students) can add questions for practice exams. A very interesting research question is to how to select the best questions among those created by the users. This could be done on the basis of difficulty, explicit feedback and implicit feedback. Based on difficulty, that could be estimated from the fraction of users answering them. Based on explicit feedback, that could be marks of users for the questions. Based on implicit feedback, that could be used by recommendation system. The recommendation system can make questions which these questions are suitable for current user (student). Well known example of recommendation system is the system used by Amazon.com. In Amazon.com, as user bought a book, the system can recommend user buy analogous books. This module is to help students revising.

In management module, administrators (teachers) will set question: choice question, true-false, essay question, manage students and check the exam scores. Teachers will modify each candidate's information and add or delete candidates. Also they will add, modify, or delete exam papers.

In online examination module, students will answer questions set by the teacher. The system will automatically generate a random test paper. The questions will be generated from the database (this questions are different from practice exams module's questions). In order to prevent candidates repeatedly refresh the papers (webpage) and make sure that the duration of the examination is recorded correctly, I designed anti-refresh function and the examination process cannot be closed, otherwise the exam results of the candidates will not normally be saved into the database. After candidates submit their papers, the system will automatically check the answer. Then candidates will be provided with a feedback on the correct answer and their scores.

## 3.2  System development tools & platforms

System platform is divided into client and server side. Due to I used browser/server (B/S) model approach to development of internet applications, so the client side will support Windows NT, Mac OS and Linux. As long as a user has a browser, he or she can use this online examination system. For clients, this is an independent system platform.

Server side used windows 2000, which is based on NT technology. It has many of advantages, including a good safety, stable running and easy managed.

My system is based on ASP.NET and SQLServer2005 development, which will eventually run on Web-server. I used following development tools, including Visual Studio.Net 2008, Macromedia Dreamweaver MX, Adobe Photoshop CS, and SQL Server 2005.

There are a lot of techniques to implement server-side programming. Such as JSP, ASP, PHP, ASP.NET. I used the current popular network programming techniques, which is ASP.NET platform supports the programming language C# as the primary means for achieving (in conjunction with JavaScript, VBScript, HTML, etc.). The efficiency of ASP.NET (as a web implementation technology) is very high. It meets the requirement of real-time examination system.

I am using Asp.net because it is a web application framework and builds on the common language runtime (CLR). It uses the server side to build powerful web applications. Compared with traditional dynamic pages technologies, the program design of Asp.net is relatively simplified and its structure is clear. However, the cost is that the page design is more difficult and the interface needs to be handled. It is difficult to use the other tools. Asp.net separates program from interface. In other words, Asp.net makes web applications like traditional windows applications. The designer has easy jobs. He can look at Datagrid / Datalist / calendar / Repeater and other web controls. You will find that these well-designed reusable controls greatly improve the efficiency of programming. Dreamweaver is a powerful and efficient tool for web page designer. I think asp.net is a major reform for programmers. It makes the large web applications much easier and in accordance with modern software engineering. In this point, web applications can be seen as the distributed expansion of the traditional C# structure.

## 3.3 Database design

My system used Microsoft SQL Server 2005 database, which is a high performance relational database management system. I built several tables which including candidate information table, questions table, exam paper table, user answer table, student's score table and so on.
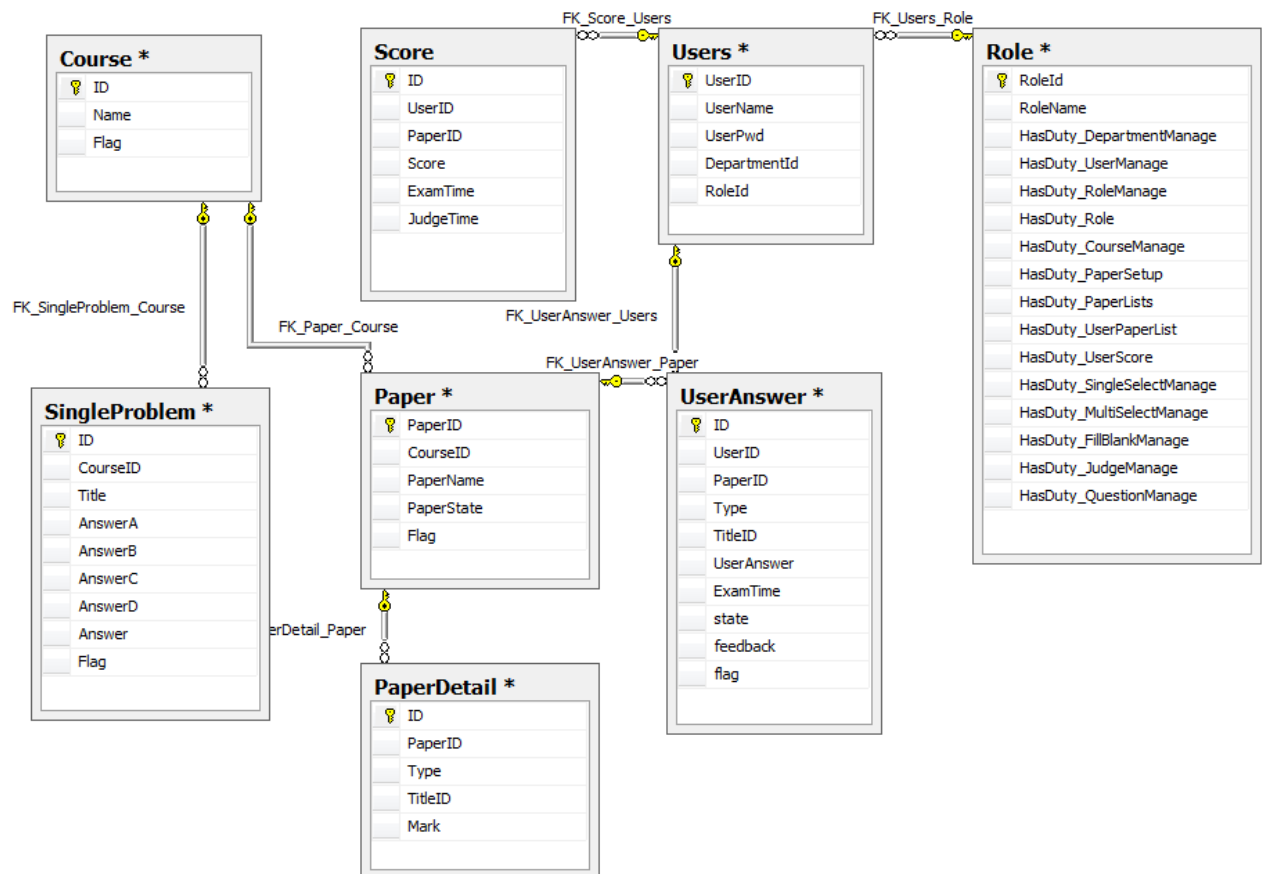


Figure 3.3-0 relationship of each table

In the relationship of each table, using *UserID* connected *Users* table, *Score* table and *UserAnswer* table together. Hence, using *UserID* can find the student's score and his/her answer sheet. The *Role* table shown different user has different duty. It used *RoleId* connect to *Users* table, so we can find a user's role in the system. Due to the page limit, I just show single choice question in the relationship of each table. Using *CourseID* connected *Course* table, *Paper* table and *question* table. Hence, we can find paper information through *CourseID* and *PaperID*.

Each table details are following:

| Column Name | Data Type | Allow Nulls |
| --- | --- | --- |
| UserID | varchar(50) | ☐ |
| UserName | varchar(50) | ☐ |
| UserPwd | varchar(64) | ☐ |
| DepartmentId | int | ☐ |
| RoleId | int | ☑ |

Figure 3.3-1 Candidate information table

| Column Name | Data Type | Allow Nulls |
| --- | --- | --- |
| ID | int | ☐ |
| CourseID | int | ☐ |
| Title | varchar(1000) | ☐ |
| AnswerA | varchar(500) | ☐ |
| AnswerB | varchar(500) | ☐ |
| AnswerC | varchar(500) | ☐ |
| AnswerD | varchar(500) | ☐ |
| Answer | varchar(10) | ☐ |
| Flag | varchar(50) | ☑ |

Figure 3.3-2 Single choice question table

The 'Flag' is used classify the question whether belongs to real exam or practice exam. The rest of question type is similar with single choice question.

| Column Name | Data Type | Allow Nulls | Column Name | Data Type | Allow Nulls |
| --- | --- | --- | --- | --- | --- |
| PaperID | int | ☐ | ID | int | ☐ |
| CourseID | int | ☐ | PaperID | int | ☐ |
| PaperName | varchar(250) | ☐ | Type | varchar(100) | ☑ |
| PaperState | bit | ☐ | TitleID | int | ☐ |
| Flag | varchar(50) | ☑ | Mark | int | ☐ |

Figure 3.3-3 Paper table 1 and 2

The left side of paper table is recorded paper name, id and paper state. The right side of paper table is recorded more details. The PaperID of right side table is a foreign key for table 1, which means both of paper tables are connected. The TitleID of right side table is a foreign key for question table. Hence, the question table and paper table are also connected.

Figure 3.3-4 User answer table

The feedback is recorded a rating of current question. This feedback is used recommendation questions.



Figure 3.3-5 score of student table

The PaperID is a foreign key for paper table. It can find exam paper though PaperID.

## 3.3.1 Using stored procedures in database

"A stored procedure is a subroutine available to applications accessing a relational database system. Stored procedures are actually stored in the database data dictionary."[24]. User can execute the stored procedures by the procedures name and parameters. Using SQL statements to write a stored procedure has the following advantages [25]:

1) Before running the stored procedure, the database parsing was carried out, which can improve the execution performance of SQL statements. That is why to execute stored procedure is faster than normal SQL statements.

2) Stored procedure can guarantee the security of data. Stored procedure declined the user access if the user does not have permission.

In order to improve the operating efficiency of the server, my system also use a variety of database stored procedures. For example, to add a user in database:

```
CREATE PROCEDURE [dbo].[Proc_UsersAdd]
    (@UserID        [varchar](20),
     @UserName      [varchar](20),
     @UserPwd       [varchar](64),
     @DepartmentId  [int],
     @RoleId        [int]
     )

AS INSERT INTO [OnLineExam].[dbo].[Users]
    ([UserID],
     [UserName],
     [UserPwd],
     [DepartmentId],
     [RoleId]
     )

VALUES
    ( @UserID,
     @UserName,
     @UserPwd,
     @DepartmentId,
     @RoleId
     )
```

Figure 3.3.1 Add user stored procedure

# 3.4  System security design

The security of online examination system is vital for the candidates. Good security is basic conditions to ensure a fair examination. The system will not be widely used without good security design. To ensure system security and confidentiality, I mainly achieved through the following:

1)  All candidates who access to the system must be authenticated while registering their computer IP address. To prevent cheating, the system forbids two or more candidates using the same account to login. In addition, system only lets candidates login once. (Under special circumstances, for example, her or his computer crashed, to remove this restriction by the administrator)

2)  In order to ensure stability of the system, system should verify the users whether exceed the maximum login size restrictions. Therefore, we need to set the maximum number of login restrictions. Over the limit, the system will forbid the follow-up user login. After entering the system, if the user does not make any action within the specified time, his/her all operations will be cancelled, and the

user must re-login if he/she wants. At the same time, the system login authentication mechanism not only achieves the user authentication, but also be able to intercept the user input a URL address in a browser, and prevent users to enter the URL address without login. This design can prevent illegal system entry.

3) System can assign different operations permissions to different users. System can forbid users viewing all examination papers (except current paper) and copy the paper.

### 3.4.1 Prevent cheating

The security requirement of online examination system is very strict. The system forbids two or more candidates using the same account to login. Hence, candidates can be avoided using a user name and password to login by multiple IP addresses at the same time. This is called Single Sign On (SSO). It can effectively prevent cheating.

To achieve SSO, it mainly uses Cache object. Cache object is usually used for caching Web applications. Basically, I saved the login information of candidates into Cache object and set the expired time of Cache being the same as Session's expired time. Hence, once the Session is expired, the Cache is also expired. Since the Cache object has already stored login information of a candidate, if the candidate logins twice or more, the system will refuse him/her.

In addition, to prevent that candidates using internet to search the answer or use copy, paste and refresh to affect exam, the ctrl, shift, F5 on the keyboard must be shielded and disconnect the outer net during the exam.

### 3.4.2 Prevent skipping login interface

Login screen is the first security barrier in my system. The illegal user (hacker) may enter a URL address and directly go into a page. After that, hacker may steal the user registration information and exam paper information. In order to prevent such situation, I have established the identity of an authentication session object to determine the legality of the user login. If someone skips the login page, the system will prompt an error message and return to the login page. For example, a candidate normally logins, then set the state of *Session* as *true:*
*Session ("student") = true;*

If the candidate hasn't login, system will set the state of *Session* as false and redirect to login page.

*If (Session ("student") == false)*

*{*

    *Response.redirect (login.aspx)*

*}*

### 3.4.3 Prevent SQL injection attack

The SQL injection attack is an attempt by a attacker to insert SQL command into the Web form and deceive the server to execute malicious commands. For example, in the login screen, if I write the code to verify user's name and password like this: *select \* from UserTable where username="&user&" and password="&pwd&"",* and a hacker input the username as "1 or 1=1" and password as the same with username, the code will change to: *select \* from UserTable where username=1 or 1=1 and password=1 or 1=1.* The system will pass this verification and hacker can access the system without real username and password. Therefore, I defined a function to filter special character.

*pulic string Consql (stringistring)*
*{*
*istring=istring.Trim();*
*istirng=String.RePlace("or", "");*
*istring=istring.RePlaee("and", "");*
*istring=istrillg.RePlaee("!", "");*
*istring=istring.RePlace("+", "");*
*istring=istring.RePlaee("—", "");*
*returnistring;*
*}*

### 3.4.4 IIS security configuration

My system used Internet Information Server (IIS). Using IIS to build sever is very convenient and easy. It has become one of the most popular Web server software. IIS relates to exam paper questions information and user information, therefore the security of IIS is particularly important. I have designed the following methods to enhance security of IIS.

1) Remove unnecessary virtual directory
   After installing the IIS, the system generates default directory in root direcroty,

including IISHelp, IISAdmin, IISSamples and MSADC. These directories are useless. They are therefore deleted directly

2) Remove the risk of IIS components

After installing the IIS, some of IIS components may lead to security threats, such as Simple Mail Transfer Protocol (SMTP) Service and Network News Transport Protocol (NNTP) Service. If we do not use the components, we should remove them as much as possible

3) Classify IIS files and Set permissions for each file

According to different types of directory, appropriate permissions should be assigned. For example, static file folders (HTML file) allow to read, refuse to write. ASP.NET scripts folders allow implementation, refuse to write and read. EXE and other executable program allow implementation, refuse to read and write.

4) Remove unnecessary application mapping

There are many applications mapping in the default IIS configuration. In addition to ASP.NET mapping, other mappings are rarely used. We should remove useless mapping using the "Internet Services Manager".

5) Protect the log security

The log is an important part of strategy to improve overall system security. By default, the IIS log is saved at %WinDir%\System32\LogFiles. The hacker should know the path, so I need to change a path and change the access permissions in which only administrator can access the log file.

Through the above security settings, we can enhance the security of server.

# Chapter 4

# System implementation and key technology

## 4.1  Public class

In the project, I created several classes to package some common methods or functions. These classes not only improve the reuse of code, but also to facilitate code management. SQL Server 2005 is a large relational database system. In ASP.NET, I used ADO.NET technology to access SQL Server. The following is several classes of ADO.NET for accessing the SQL Server:

SqlConnection: this class is used for establishing the connection to SQL Server, which open database connection.
SqlCommand: this class is used for executing the SQL statement or stored procedure.
SqlDataAdapter: this class is used for padding the dataset or update database.
SqlDataReader: this class is used for reading data from database.
SqlParameter: this class is used for transmitting parameters to store procedure.

I created a database class, which including following functions:
The *Open ()* function is used for connecting to the database and open the connection.
The *Close ()* function is used for disconnecting to the database and close the connection.
The *GetRecord(string XSqlString)* function is used for getting query resulting from database.
The *RunProc(string ProcName)* function is used for executing a stored procedure.
The *RunProc(string ProcName, SqlParameter[] Params)* function is used for

executing a stored procedure with a parameter of the stored procedure.
The *GetDataSet(string ProcName)* is used for executing a stored procedure and svaing the result into a dataset.


## 4.2  Login module implementation

To prevent the illegal system access, login module can verify the legitimacy of the user. A user must enter his/her username and password. In order to enhance security, this system used verification code technology. The system can prevent the robot program to login system repeatedly through the randomly generated verification code. Users enter a user name and password, click "Login" button to log on. In the process, the system verifies the username, password and verification code. According to the identity of the user, system will show different page to user.



Figure 4.2 Login module

Login module used following control component:

TextBox: text box component for user input text.

Button: the 'Login' button submits the username, password and verification code to server.

RequiredFieldValidator: this component checks whether the text box is null. If the user hasn't entered username or password, it will show error message.

Verification code: this component creates four digital numbers randomly and generates a figure. It can prevent hacker using robot program to login system repeatedly.

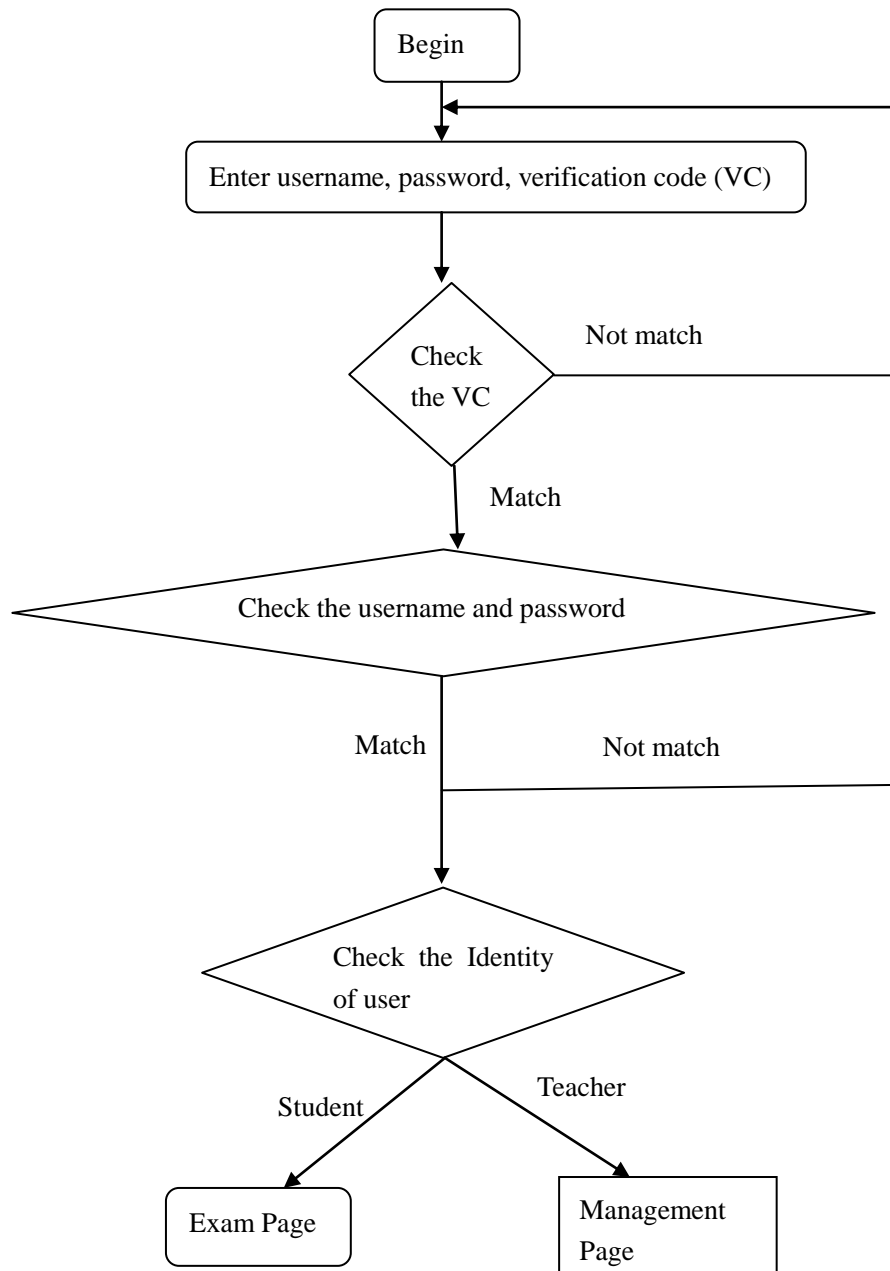The flow char of login module is as follows:

Figure 4.2-1 flow of login module

## 4.3  Register module implementation

In the real exam, only the teacher can register a user to a student. In the practice exam, student can register a user for exercises. Student should enter his/her ID, name, password, and choose department. To prevent student using the same ID to register, the system can check the ID in the database. If the same ID is found, the system will show an error message. After that, system saved all student information into the database.

Figure 4.3 register a new user.

## 4.4  User management

To take part in the real exam, student needs a user id and password, which need teacher to add a new user for a student (Figure 4.4-1). The teacher also can modify and delete the user. If the student forgot his/her password, teach can reset his/her password and send a new password to the student (Figure 4.4-2).



Figure 4.4-1 Add a user

Figure 4.4-2 User Management


# 4.5  Management of question bank

After a teacher logins the system, he can manage the question bank, including to add question into the database and to modify the question information. There are five types of question in this system: single choice, multiple choices, true-false, fill-in-blank and essay questions. The teacher should choose exam course name and enter question title, answer of each selection and correct answer. The correct answer is used for automatic marking paper. For example: add single choice question (Figure 4.5-1).

Figure 4.5-1 Add single choice questions into the database

Teacher can manage the question. Click 'Details' of each question, they can modify question information. Click 'Delete' for deleting the question. The following figure is management of single choice question.


Figure 4.5-2 Manage single choice question

## 4.6 Automatically generate test paper based on genetic algorithm

Genetic algorithm generally uses a simple code to represent complex structures, which uses binary string with fixed-length to encode. A solution of the question corresponds to a binary string of length *k*. Start from the initial population, calculate the fitness of each individual in current population, use crossover and mutation to produce the next generation of population. In this way, a new generation is made continuously until the terminal condition is satisfied.
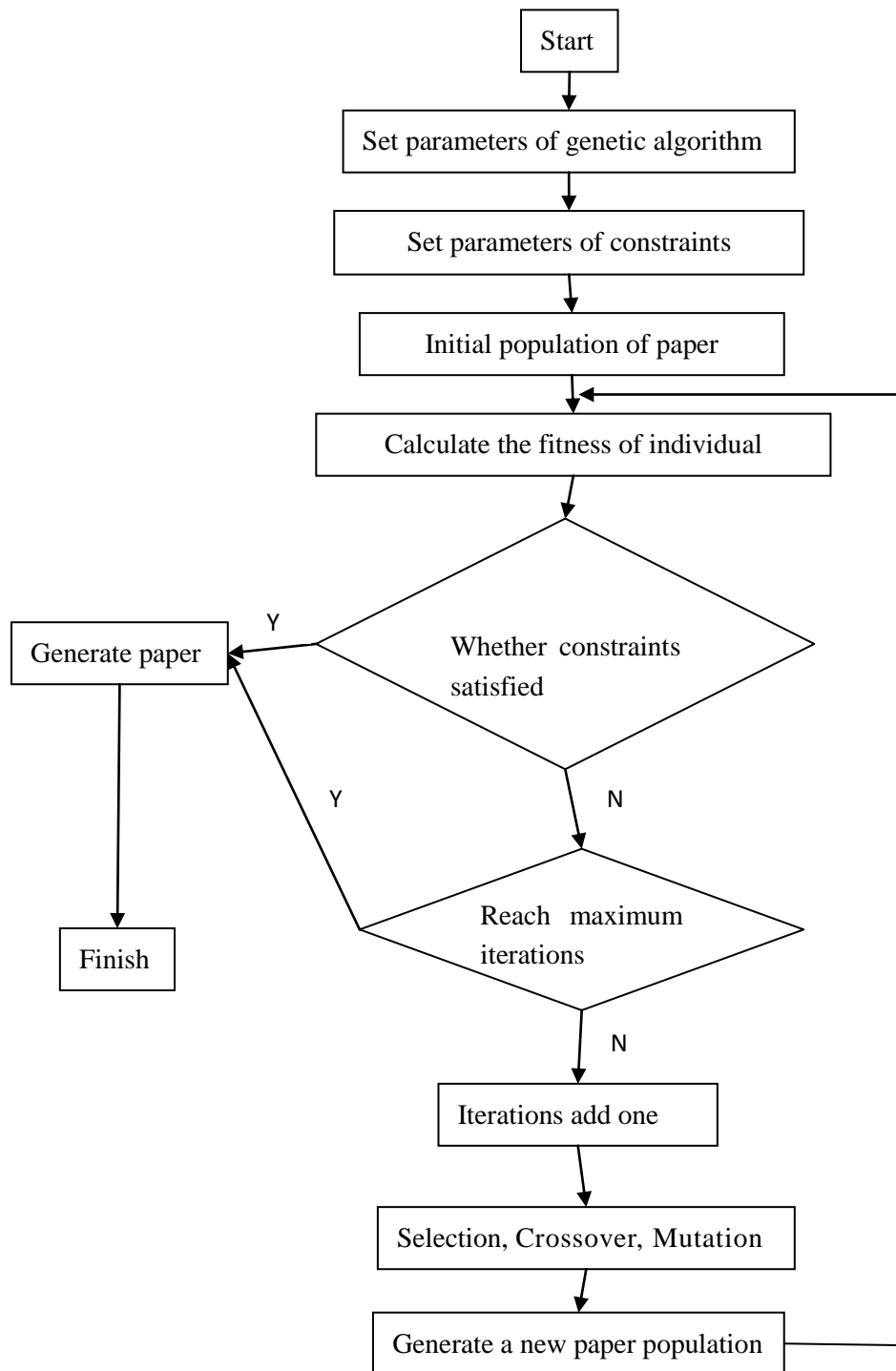


Figure 4.6-1, genetic algorithm diagram of generating papers. [18]

Specific steps of setting test paper are as follows:

1) Using binary string coding to represent questions

The length of coding is decided by the number of questions in question bank. Assuming the question bank has *m* questions $X_1, X_2, X_3.......X_m$, the questions of test paper are selected from $X_1, X_2, X_3.......X_m$. If $X_i$ *(i=1, 2, 3 ... m)* = 1, the question is selected. If $X_i$ *(i=1, 2, 3 ... m)* = 0, the question is not selected. For example, assuming a test paper includes 10 questions, there are ten '1' in $X_1, X_2, X_3.......X_m$. therefore a piece of paper can represent a binary string of '010110110101110…000'.

2) Initial population

In order to make the system keep equal opportunities in the initial search for each population, we usually generate a string population by random initialization. In string population, length of string is the same. The size of initial population is usually decided from the specific problems, or given by experience or experiment [20]. When system generates test paper, the size of population would be set *n*. Each individual randomly selects questions by function of random (m).'m' is the number of Questions in question bank.

Initialization process is as follows:

*for i=1 to population_size do*

*random (m)*

*end*

3) Calculate the fitness of each individual in current population

To calculate the fitness of each individual, a function of fitness is required. In this report, the function of fitness is:

$$F = \frac{1}{\left(1+\sum_{i-1}^{12} k_i \, |e_i| \right)}$$

*i* = 1, 2, 3…n.     $k_i > 0$

The $e_i$ denotes the error of $i^{th}$ paper. The $k_i$ denotes weight coefficient. If the error of test paper of individual constraints is small, the value of fitness is greater, which means the individual (question) is close to targets of test paper.

4) Selection

"Selection is the stage of a genetic algorithm in which individual genomes are chosen from a population for later breeding (recombination or crossover)." (Wikipedia Selection genetic algorithm). The purpose of selection is to select good individuals from the current population, so that they have the opportunity to breed the next generation. Selection is guided by the principle of 'survival of the fittest'.

Generally, we use 'Roulette wheel selection' (figure 3) to choose individual from a population. The roulette wheel selection is similar to gambling of roulette. According

to the individual fitness, roulette size is divided into different proportion of the area. Individuals with high fitness occupy a large proportion of the roulette wheel, the probability that individuals are selected is larger.
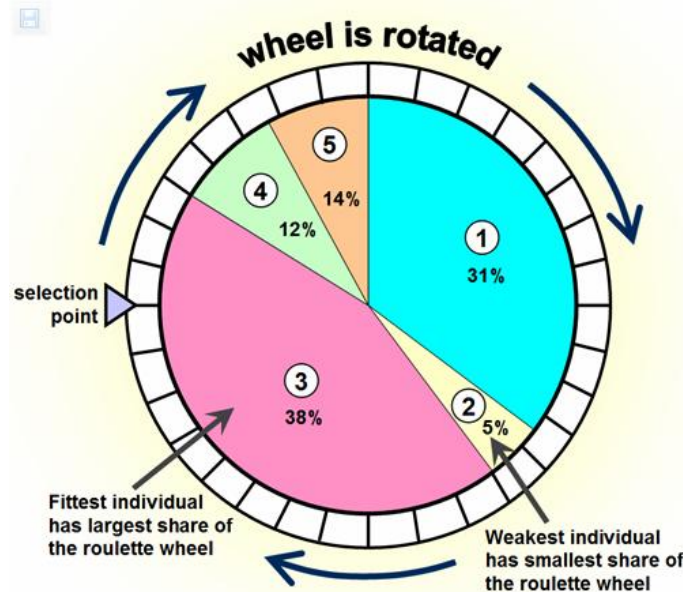


Figure 4.6-2 Rulette wheel selection
Source: http://www.edc.ncl.ac.uk/highlight/rhjanuary2007g02.php/

   5)  Crossover

"Crossover used to vary the programming of a chromosome or chromosomes from one generation to the next. It is analogous to reproduction and biological crossover." (Wikipedia Crossover genetic algorithm). First, two individuals are randomly selected from population, and then a crossover point is randomly selected. For example, there are two individuals A1 and A2. The crossover point is randomly selected by 7. The length of string is 10.

  A1 = 1000011111
  A2 = 1111111011
  After crossover, got new A1' and A2':
  A1'＝1000011011
  A2'＝1111111111

   6)  Mutation

The operation of mutation changes the value of the string from 0 to 1, or from 1 to 0. However, not all the bits can be changed. The probability of mutation of each bit is $P_m$. $P_m$ is specified a value which is between 0 and 1. In each bit of string, the mutation is independent. That is, one mutation does not affect another mutation. The effect of mutation is to introduce new genetic material. For example, if the value of a string is 0, the crossover can never produce the value of 1. Only by mutation, the value of string can change.

   7)  Termination

Genetic algorithm is a kind of iterative search process. This process is repeated until a termination condition has been reached. There are two kinds of methods of termination. The first method sets fixed number of iteration. At the beginning, the number should be small, such as 100 times. Then according to the situation (e.g. the constraints are not sufficiently satisfied) gradually increase, can reach thousands of times. Then the number of iteration can gradually increase, even reach thousands of times. The second method is that termination is decided by fitness of individual. Once the fitness of individual does not change or changes little, the process of iteration will be terminated. I combined the above two methods in my project.

# 4.7 Using Slope One Algorithm to Recommend Questions

In the section 2.5.6, we introduced some conception of slope one algorithm, now I want to introduce how to implement slope one algorithm in my system.
Basically, Using slope one algorithm need following data:
1) Some several users
2) Some several questions
3) Some users' rating of questions

Slope one algorithm to solve the problem that is recommends questions which haven't rated to users according to some several ratings.
To implement slope one recommendation system, we need following steps:
1) Loading the user's rating from database and save the rating in memory
2) Calculated rating difference of any two questions
3) Input user's several rating of question, system predicts the possible rating of the rest questions.
4) Sorting the prediction, and show the Top Questions.

Step one: Loading the user's rating from database. Assume there are three users and four questions. User rating for each question is following:

| Ratings | User1 | User2 | User3 |
|---------|-------|-------|-------|
| Question 1 | 5 | 4 | 4 |
| Question 2 | 4 | 5 | 4 |
| Question 3 | 4 | 3 | N/A |
| Question 4 | N/A | 5 | 5 |

I need to calculate the rating difference of any two questions. So I got the following

matrix:

|  | Item1 | Item2 | Item3 | Item4 |
|---|---|---|---|---|
| Item1 | N/A | 0/3 | 2/2 | -2/2 |
| Item2 | 0/3 | N/A | 2/2 | -1/2 |
| Item3 | -2/2 | -2/2 | N/A | -2/1 |
| Item4 | 2/2 | 1/2 | 2/1 | N/A |

Considering the weight slope one, I need to record the number of users having rated both items. Firstly, I defined a class to store rating (c#):

```csharp
public class Rating
{
    public float Value { get; set; }
    public int Freq { get; set; }

    public float AverageValue
    {
        get { return Value / Freq; }
    }
}
```

Secondly, I used "Dictionary" (C# Generic collection) to store the rating difference (the matrix).

```csharp
public class RatingDifferenceCollection : Dictionary<string, Rating>
{
    private string GetKey(int Item1Id, int Item2Id)
    {
        return (Item1Id < Item2Id) ? Item1Id + "/" + Item2Id : Item2Id + "/" + Item1Id;
    }

    public bool Contains(int Item1Id, int Item2Id)
    {
        return this.Keys.Contains<string>(GetKey(Item1Id, Item2Id));
    }

    public Rating this[int Item1Id, int Item2Id]
    {
        get
        {
            return this[this.GetKey(Item1Id, Item2Id)];
        }
        set { this[this.GetKey(Item1Id, Item2Id)] = value; }
    }
}
```

The Dictionary generic class provides a mapping from a set of keys to a set of values. Each addition to the dictionary consists of a value and its associated key. Retrieving a value by using its key is very fast, close to O (1), because the Dictionary class is implemented as a hash table. [26]

Thirdly, I need to predict the possible rating of the rest questions, it need two steps.
1) I wrote a class called SlopeOne.

```
public class Slopeone
{
    public RatingDifferenceCollection _DiffMarix = new RatingDifferenceCollection();
    public HashSet<int> numOfItems = new HashSet<int>();

    public void AddUserRatings(IDictionary<int, float> userRatings)[...]


    public IDictionary<int, float> Predict(IDictionary<int, float> userRatings)[...]
}
```

I need instantiation the RatingDifferenceCollection (_DiffMarix) to store the rating difference (the matrix), and then declared HashSet to store the number of questions. The function of AddUserRating can accept a record of user's rating. There are double loop in the AddUserRating function. Surrounding loop traverses all rating of questions, inner-loop traverses again and calculates the rating difference of a pair of questions. Since I used weight slope one, I need calculate the weight. In program, I used Freq plus 1 if these pair of questions appears once. The following code is from inner-loop:

```
Rating ratingDiff;
if (_DiffMarix.Contains(item1Id, item2Id))
{
    ratingDiff = _DiffMarix[item1Id, item2Id];
}
else
{
    ratingDiff = new Rating();
    _DiffMarix[item1Id, item2Id] = ratingDiff;
}

ratingDiff.Value += item1Rating - item2Rating;
ratingDiff.Freq += 1;
```

After every users called the function of AddUserRating, the system built a matrix, but my matrix is a table format.

|  | Rating Difference | Freq |
|---|---|---|
| Item1-2 | 0 | 3 |
| Item1-3 | 1 | 2 |
| Item2-1 | 0 | 3 |
| Item2-3 | 1 | 2 |
| Item3-1 | -1 | 2 |
| Item3-2 | -1 | 2 |
| Item1-4 | -1 | 2 |
| Item2-4 | -0.5 | 2 |
| Item3-4 | -2 | 1 |
| Item4-1 | 1 | 2 |
| Item4-2 | 0.5 | 2 |
| Item4-3 | 2 | 1 |

2) Input a user's rating, predict the rating of the rest of questions:

```
public IDictionary<int, float> Predict(IDictionary<int, float> userRatings)...
```

In function of Predict, there are also double loop. Surrounding loop traverses all rating of questions in the HashSet, and inner-loop traverses current user's ratings which from the parameter. Using this user's rating and the matrix, system can calculate the rating of the rest of questions:

```
Rating diff = _DiffMarix[itemId, inputItemId];
itemRating.Value += diff.Freq * (userRating.Value + diff.AverageValue * ((itemId < inputItemId) ? 1 : -1));
itemRating.Freq += diff.Freq;
```

Observing the matrix, we can see the value of diagonal in the matrix never has a value and the value of below the diagonal is symmetrical with the value of above the diagonal. The value of below the diagonal is equal the value of above the diagonal multiplied -1. Hence, I add "(itemId < inputItemId) ? 1 : -1)" to save the store space.

Finally, sort the prediction, and show the Top Questions to users.

# Chapter 5

# Testing

## 5.1 System function testing

The function testing belongs to black-box testing. It is used for detecting errors of each function of system. Normally, black-box testing is used mainly in order to find the following types of errors: [27]

1) Whether the function is incorrect or missing.
2) Whether the data structure is incorrect or external information (e.g. data files) accessed error.
3) Whether the performance is satisfied requirements.
4) Whether the Initialization or termination is incorrect.

In testing, the system looks like a closed black-box. Without considering the internal structure of the system, we test the system interface. It checks the function of system whether is in accordance with the specification. It also checks the system whether receive input data to generate the correct output data. Black-box testing focuses on the external structure of system and don't consider the internal structure.

Using the black-box testing has following advantages:
It is simple. Tester does not need to understand the system code and how to implementation.
From the user point of view, tester can easily know the user will be used function of

the system and also know the possible encountered problem.

The following is testing user login module.

| Function | User authentication |
|---|---|
| Test Purpose | Verify that the input information to allow the legal login, or to stop the illegal login |
| Test data (correct data) | Username: user123, password:123user |

**Table 5.1 Test case – login module**

| Number of testing | Input data | Output result | Expect result |
|---|---|---|---|
| 1 | Username = 'user123'<br>Password = '123user'<br>Verification code = '2076' | Successfully login and jump to exam screen | Successfully login and jump to exam screen |
| 2 | Username = 'user123'<br>Password = '123user'<br>Verification code: deliberately input a incorrect number, e.g. 2343 | Show a message on screen: The verification code is incorrect. | Show an error message about verification code. |
| 3 | Username = 'user123'<br>Password = ' '<br>Verification code = '5749' | Show a message on screen: the password is not allowed to a null. | Show an error message about the password |
| 4 | Username = ' '<br>Password = ''<br>Verification code = '5930' | Show two messages on screen: The username is not allowed to a null. The password is not allowed to a null. | Show two messages on screen: username and password are not allowed to null. |
| 5 | Username = 'user123'<br>Password = 'werer'<br>Verification code = '1049' | Show a message on screen: your username or password is incorrect. | Show a message on screen: your password is incorrect. |

**Table 5.2 Test progress – login module**

The generating exam paper testing is as follows:

| Function | User authentication |
|---|---|
| Test Purpose | Verify the result of generating paper |
| Test data | 60 single choice questions from driver course |

**Table 5.3 Test cases – generate exam paper**

To generate an exam paper, the teacher needs to enter number of questions and each question's credit. So let NQ is number of questions. Let QC is credit of question.

| Number of testing | Input data | Output result | Expect result |
|---|---|---|---|
| 1 | NQ = 10<br>QC = 5 | Question ID: 23, 3, 11, 2, 4, 28, 45, 7, 34, 29<br>And generate the paper successful. | The question ID is randomly selected and successful generate exam paper. |
| 2 | NQ = -1<br>QC = 5 | Show an error message on screen: you enter an incorrect value. Try again. | Show an error message and doesn't generate the exam paper |
| 3 | NQ = 10<br>QC = 0 | Show an error message on screen: the credit of question should not less than 1. Try again. | Show an error message and doesn't generate the exam paper. |
| 4 | NQ = 100<br>QC = 1 | Question ID: all of questions (60) and generate the exam paper. | Show all of questions, because there is no enough questions to select. |

**Table 5.4 Test progress – generate exam paper**

Testing recommendation of questions:
Dataset: 7 users and 12 questions. The rating of each question is as follows:

|  | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 | Q12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| User1 | 1 | 0 | 1 | 1 | -1 | 0 | 0 | 0 | 1 | -1 | 1 | -1 |
| User2 | -1 | -1 | 1 | 1 | 1 | 1 | 1 | -1 | 0 | 0 | 1 | -1 |
| User3 | 0 | -1 | 0 | -1 | 1 | 1 | -1 | 1 | 0 | 0 | -1 | -1 |
| User4 | 1 | 1 | 0 | 1 | 1 | 1 | -1 | 1 | 1 | 0 | 1 | -1 |
| User5 | 1 | -1 | 0 | 1 | -1 | -1 | 1 | 1 | 1 | 0 | -1 | 1 |
| User6 | 1 | -1 | 1 | 1 | -1 | -1 | 1 | 1 | -1 | 1 | 0 | 0 |

| User7 | | -1 | | | | 1 | 0 | | -1 | | 0 | |

**Table 5.5 User's rating for each question**

User 1 to user 6 have done all of questions and given them ratings of each question. User 7 has done 5 questions and given them ratings of each question. When user 7 wants to do practice exam again, the system will recommend question 4 to user 7 (if he need one question). Because of the prediction rating of question 4 is 0.4.

| | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 | Q12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| User1 | 1 | 0 | 1 | 1 | -1 | 0 | 0 | 0 | 1 | -1 | 1 | -1 |
| User2 | -1 | -1 | 1 | 1 | 1 | 1 | 1 | -1 | 0 | 0 | 1 | -1 |
| User3 | 0 | -1 | 0 | -1 | 1 | 1 | -1 | 1 | 0 | 0 | -1 | -1 |
| User4 | 1 | 1 | 0 | 1 | 1 | 1 | -1 | 1 | 1 | 0 | 1 | -1 |
| User5 | 1 | -1 | 0 | 1 | -1 | -1 | 1 | 1 | 1 | 0 | -1 | 1 |
| User6 | 1 | -1 | 1 | 1 | -1 | -1 | 1 | 1 | -1 | 1 | 0 | 0 |
| User7 | 0.23 | -1 | 0.23 | 0.4 | -0.26 | 1 | 0 | 0.23 | -1 | -0.26 | 0 | -0.76 |

**Table 5.6 Predict rating for the rest questions**

# Chapter 6

## Conclusions & further work

This thesis for the first time investigates how recommendation algorithm can be employed to improve examination systems. This thesis firstly introduced various kinds of existing online exam systems and software development models. Secondly, it introduced some algorithms for generating exam paper and some algorithms for recommendation questions. Through the comparison of recommendation algorithm, I chose the slope one algorithm to implement online exam system.

This system used B/S software model to development. Database used SQL Server 2005 and development environment used Visual Studio 2008, which can be made the system more quickly respond. This system developed in this project is divided two parts, which are real exam and practice exam. In the real exam, teacher sets an exam and student takes part in the exam. In the practice exam, student can set an exam and take part in the exam. If the student is first time to set an exam, the system will randomly choose questions for students. If the student has already set the exam, the system will recommend his/her interested questions. In the same time, I used a variety of security policies to ensure the security of the system.

Finally, I tested the system by the black-box testing. According to the testing result, it is found that the system is stable, and achieves objectives of the proposal. The system supports different courses, different kinds of questions and successfully generates exam paper or recommendation questions.

However, in the actual operation, some shortcoming s have grdually arisen in the system. Thest shortcoming s include:

1) The exam questions presented only in the form of plain text. I will continue my research and make it possible for the system to show languages and images in the form of multimedia.
2) I need to strengthen the function of score statistics and the function of papers analysis in the future work.
3) This system cannot mark essay question automatically. In the future work, I will apply artificial intelligence technology to achieve making essay question.

# Reference:

1. Sarwar B., Karypis G., Konstan J., and RiedlJ..*Item-based Collaborative Filtering Recommendation Algorithms.*10th international WWW conference, Hong Kong, May 2001.
2. Vucetic, Obradovic.*A Regression-Based Approach for Scaling-Up Personalized Recommender Systems in E-Commerce 2000*
3. *Breese J，Hecherman D，Kadie C．Empirical analysis ofpredictive algorithms for collaborative filtering．In：Proceedings ofthe 14th Conference on Uncertainty in Artificial Intelligence(UAI'98)．1998．43-52*
4. Adomavicius, G, Tuzhilin, A., *Toward the next generation of recommender systems: a survey of the state-of-the –art and possible extensions,* IEEE transactions on knowledge and data engineering 17, 2005
5. Sarwar B., Karypis G., Konstan J., and Riedl J. *Analysis of recommendation algorithms for e-commerce,* proceeding of the 2nd ACM conference on Electronic commerce, 2000
6. Jiru Ying, Lixin Hu, *Creating and Realizing a Model of Selecting Items Randomly from Examination Items Library*. Computer application.2000 p46-47.
7. June Jamrich Parsons，Dan Oja.*New Perspectives on Computer Concepts, Ninth Edition, Introductory.*Course Technology; 9 edition (May 8, 2006)
8. Ming Zhou, Shudong Sun, *Principle and Application of genetic algorithms.*BeijingNational Defence Industry Press. 2002
9. *HaitaoRen. Genetic algorithm used in Examination system.* Journal of chifeng university. 2009.Volume25(8)
10. Bozman, Jean; Gillen, Al; Kolodgy, Charles; Kusnetzky, Dan; Perry, Randy; &Shiang, David (October 2002). *Windows 2000 Versus Linux in Enterprise Computing: An assessment of business value for selected workloads.* IDC, sponsored by Microsoft Corporation.White paper.
11. Andrew J. Brust,Stephen Forte*(2006).Programming Microsoft SQL Server 2005.*Microsoft Press
12. Lance Delano, Rajesh George et al. (2005). *Wrox's SQL Server 2005 Express Edition Starter Kit (Programmer to Programmer).* Microsoft Press
13. Joy Mundy, *Using Partitions in a Microsoft SQL Server 2000 Data Warehouse*, http://msdn.microsoft.com/en-us/library/aa902650(SQL.80).aspx (access 01/05/2010)
14. Hai Pang. *The online examination system based on ASP.NET Analysis and Implementation.*Science and technology information.2008，Volume 24
15. Liang Zhang, Guohua Zhan, Zhenming Yuan. *A Long-distance Test System for Computer Operation Skill Based on Intemet Technology*.Computer applications. 2001 Volume 21(7)
16. Yang Xiao, Xiao Wang, *An algorithm for setting online test papers*. Journal of Beijing University of Chemical technology. 2006. Volume 33(4)

17. Werry C. *The work of education in the age of ecollege* [J ] .Computers and Composition , 2002 , 19 : 127 - 149.

18. Holland J．H，Genetic Algorithms and Classifier Systems：Foundations and FutureDirections，Genetic Algorithmsand Their Applications，Proc ofthe 2rid Int On GeneticAlgorithms，1987，82～89

19. J C．Ports，T．Giddens，B．Yadav Surya，The Development and Evaluation ofanIrnprvedGenetic Algorithm Based on Migration and Artificial Selection[J]，IEEE TransactionsOn Systems，Man and Cybernetics，1994，24(1)：73～86

20. Qing Yang, Research on Problem of Item Bank Auto-Composing Test Paper Based on Genetic Algorithm. Journal of JiNan University Science and Technology. 2004. Volume 18 (3)

21. Daniel Lemire, Anna Maclachlan, Slope One Predictors for Online Rating-Based Collaborative Filtering, In SIAM Data Mining (SDM'05), Newport Beach, California, April 21-23, 2005.

22. Slope one, Wikipedia, http://en.wikipedia.org/wiki/Slope_One (access 31/08/2010)

23. DONG LI, XING Chunxiao, WANG Kehong, Collaborative filtering algorithm evaluation for various datasets. Journal of Tsinghua University Science and Technology, 2009, Vol 49, No.4

24. Stored Procedures, http://en.wikipedia.org/wiki/Stored_procedures (access 05/09/2010)

25. Liu Weiguo, Ya Hui, Database technique and application: SQL Server, Beijing Tsinghua University, 2007

26. Dictionary class, http://msdn.microsoft.com/en-us/library/xfhwa508.aspx (access 05/09/2010)

27. Rajib Mall. Fundamentals of software engineering second edition. Prentice Hall, 2002