**《计算机网络》题库管理系统**

1. **绪论**

**1.1课题研究的目的和意义**

从古至今，考试都是教学过程中很重要的一个环节，一方面可以检查学生的学习情况，另一方面也是教师对教学成果的检查。考试有笔试、口试上机操作考核等，无论哪一种形式试卷都是考试的重要载体。传统的考试是由教师人工完成，在出卷的过程中需要教师根据多年的教学经验，依据大纲要求，查阅大量的书籍，反复推敲才能完成一份标准的试卷。在这一过程中试卷的命题内容、试题难度等由于各种因素的干扰带有较大的主观性，不够科学与客观。且教师人工出卷是一个十分耗时耗力的工程，会给教师带来很大的工作压力。

随着计算机科学和网络技术的迅速发展和广泛应用，计算机技术已进入现代教育领域并形成一门新兴的边缘科学——计算机辅助教学 CAI（Computer Aided Instruction）。利用计算机建立试题库，实现计算机自动选题组卷，是 CAI 工程的重要组成部分，是实现考试规范化、科学化的重要措施，更是实现考教分离的一个重要手段[1]。，这将使得试卷的管理与题库的管理变得方便快捷，并且会提高出题的效率减轻教师的工作压力。另一方面将使得试卷的管理走向自动化与规范化

**1.2国内外研究现状**

计算计辅助考试最早于美国诞生，历经多年已经形成了许多成熟的应用系统，比如美国政府举办的TOFEL考试，该考试利用计算机与网络技术，提高了考试的客观性与科学性也减少了美国政府在该项目上的投资。而在我国计算机等级考试等一些考试就是基于计算机技术从题库中抽取试题进行考试的。

目前的题库系统有三种模式，第一种是将试题录入题库在组卷时由人工抽取。第二种是将已经命题完毕的试卷存储到题库中，在考试时随机抽取一套进行考试。第三种是将试题录入题库根据某种组卷策略自动组卷。第一二种模式本质上还是没有摆脱传统的组卷模式，第三种模式能基本满足自动组卷的要求。其中的组卷策略是该种模式的关键与核心。常用的组卷策略有优先权算法、随机抽取算法、误差补偿算法、团溯试探法及遗传算法等。（列举各种算法的有缺点）在选择组卷算法时要考虑到组卷的用时、组卷的科学合理与客观性，比较几种组卷算法，遗传算法是较为合适的一种。

**1.3主要研究内容**

根据已有的相关课题的研究和发展趋势的分析，本系统采用JavaEE技术与遗传算法开发了一个基于B/S模式的《计算机网络》题库管理系统（以下简称题库管理系统）。该题库管理系统是为《计算机网络》定制研发的，结合《计算机网络》的学科特点与网络技术设计实现而成。该系统不受时间与空间的限制，教师与学生可随时随地通过Web浏览器进行身份验证，登录题库管理系统进行相关操作。

1. **关键技术介绍**

**2.1 开发架构介绍**

C/S结构与B/S结构是目前比较流行的两种系统体系结构，CS即Client/Server（客户机/服务器）结构，C/S结构在技术上很成熟，它的主要特点是交互性强、具有安全的存取模式、网络通信量低、响应速度快、利于处理大量数据。但是该结构的程序是针对性开发，变更不够灵活，维护和管理的难度较大。通常只局限于小型局域网，不利于扩展。并且，由于该结构的每台客户机都需要安装相应的客户端程序，分布功能弱且兼容性差，不能实现快速部署安装和配置，因此缺少通用性，具有较大的局限性。要求具有一定专业水准的技术人员去完成。

BS即Browser/Server（浏览器/服务器）结构，就是只安装维护一个服务器（Server），而客户端采用浏览器（Browse）运行软件。B/S结构应用程序相对于传统的C/S结构应用程序是一个非常大的进步。基于 B/S结构用户无需安装相应的客户端，只需要通过WWW浏览器便可以访问Internet上的资源，且后期的维护只需要修改前端页面便可以实现版本的更新与维护。

基于以上的对比B/S结构可以异地浏览且该结构下[信息采集](https://wenwen.sogou.com/s/?w=%E4%BF%A1%E6%81%AF%E9%87%87%E9%9B%86&ch=ww.xqy.chain" \t "https://www.cnblogs.com/yan1/p/_blank)十分灵活性。任何时间、任何地点、任何系统，只要可以使用浏览器上网，就可以使用B/S系统的终端。这也符合题库管理系统的要求。

**2.2 开发技术平台介绍**

目前用于B/S结构开发较为流行的平台有微软公司的Windows .NET平台和SUN公司的J2EE平台两种，题库管理系统采用的是J2EE平台。

第二代Java平台一共有三个版本（标准版J2SE、企业版J2EE、微型版J2ME）。J2EE在J2SE基础上构建，包含了J2SE中的所有类，帮助开发和部署可移植、健壮、可伸缩且安全的服务器端java应用程序。

J2EE中的13种技术规范

1、JDBC（Java Database Connectivity）

以统一的方式访问各种数据库的API

2、JNDI（Java Naming Directary Interface）

Java命名和目录接口

1. EJB（Enterprise Java Beans）

  提供一个架构来开发和配置到客户端的分布式商业逻辑。EJB规范定义了EJB组件应该如何及何时与他们的容器交互。

4、RMI（Remote Method Invoke）

远程方法调用

5、Java IDL/CORBA(通用对象请求代理架构是软件构建的一个标准 )

6、JSP（Java Server Pages）

JSP页面由[HTML](https://baike.baidu.com/item/HTML" \t "https://baike.baidu.com/item/j2ee/_blank)（[标准通用标记语言](https://baike.baidu.com/item/%E6%A0%87%E5%87%86%E9%80%9A%E7%94%A8%E6%A0%87%E8%AE%B0%E8%AF%AD%E8%A8%80" \t "https://baike.baidu.com/item/j2ee/_blank)下的一个应用）代码和嵌入其中的Java代码组成.服务器在

页面被客户端所请求以后对这些Java代码进行处理,然后将生成的HTML页面返回给客户端浏览器.

7、Java Servlet

Servlet 是一种小型的Java程序,它扩展了web服务器的功能.作为一种服务器的应用,当被请求时开始执行,这和CGI Perl脚本很相似.Servlet提供的功能大多和JSP类似,不过实现的方式不同.JSP通常是大多数的HTML代码中嵌入少量的Java代码,而servlet全部由java写成并且生成HTML.

8、XML

[XML](https://baike.baidu.com/item/XML" \t "https://baike.baidu.com/item/j2ee/_blank)（[标准通用标记语言](https://baike.baidu.com/item/%E6%A0%87%E5%87%86%E9%80%9A%E7%94%A8%E6%A0%87%E8%AE%B0%E8%AF%AD%E8%A8%80" \t "https://baike.baidu.com/item/j2ee/_blank)的子集）是一种可以用来定其它标记语言的语言.它被用来在不同的商务过程中共享数据.XML的发展和java是相互独立的,但是,它和java具有的相同目标是平台独立性.

9、JMS

MS是用于和面向对象消息的中间件相互通信的应用程序接口.它既支持点对点的域,又支持发布/订阅类型的域,并且提供了下列类型的支持:消息传递,事务型消息的传递,一致性消息和具有持久性的订阅者支持.JMS还提供了另一种方式来对新系统和旧后台系统相互集成.

10、JTA

[JTA](https://baike.baidu.com/item/JTA" \t "https://baike.baidu.com/item/j2ee/_blank)定义了一种标准API,应用程序由此可以访问各种事务监控.

11、JTS

JTS是CORBA OTS事务监控的基本实现.JTS规定了事务管理的实现方法.该事务管理器是在高层支持java Transaction API规范,并且在较低层次实现OMG OTS specification 和Java对象.JTS事务管理器为应用程序服务器,资源管理器,独立的应用以及通信资源管理器提供了事务服务.

12、JavaMail

JavaMail是用于存取邮件服务器的API,它提供了一套邮件服务器的抽象类.不仅支持SMTP服务器,也支持IMAP服务器.

13、JAF(JavaBeans ivation Framework)

JAVA安全认证框架。提供一些Act安全控制方面的框架。让开发者通过各种部署和自定义实现自己的个性安全控制策略。

**2.3 后台数据库介绍**

数据库（Database）是按照数据结构来组织、存储和管理数据的仓库，每个数据库都有一个或多个不同的API接口用于创建，访问，管理，搜索和复制所保存的数据。

我们也可以将数据存储在文件中，但是在文件中读写数据速度相对较慢，所以现在我们使用关系型数据库管理系统（RDBMS）来存储和管理的大数据量。

所谓的关系型数据库，是建立在关系模型基础上的数据库，借助于集合代数等数学概念和方法来处理数据库中的数据。

RDBMS即关系数据库管理系统(Relational Database Management System)的特点：

1.数据以表格的形式出现

2.每行为各种记录名称

3.每列为记录名称所对应的数据域

4.许多的行和列组成一张表单

5.若干的表单组成database

目前较为流行的关系型数据库有Oracle、SQL Server、Mysql等。

* Oracle是一个重量级的数据库，其运行需要消耗许多内存，并且它是一个付费的数据库。
* SQL Server是由微软开发的数据库管理系统，它只能在Windows上运行，操作系统的系统稳定性对数据库十分重要。并行实施和共存模型并不成熟，很难处理日益增多的用户数和数据卷，伸缩性有限。
* MySQL是一个开源的SQL数据库管理系统，它由MySQL AB开发、发布和支持。它既可以满足系统的开发需求，也不消耗过多的内存，并且由于MySQL是开源数据库可以降低开发成本。

1. **《计算机网络》题库管理系统需求分析**

在教学工作中，由教师担任试题库管理，而题库的维护和备份以及用户的管理由系统管理人员完成。

试题的管理包括试题的录入以及更改删除等，并且教师根据自己的组卷需求从题库中选择相应的试题组成合适的试卷。学生可以阅览试题并下载试卷。

根据以上分析，题库管理系统的需求分析如下：

**3.1题库管理系统需求分析**

**3.1.1 题库管理系统业务流程图**

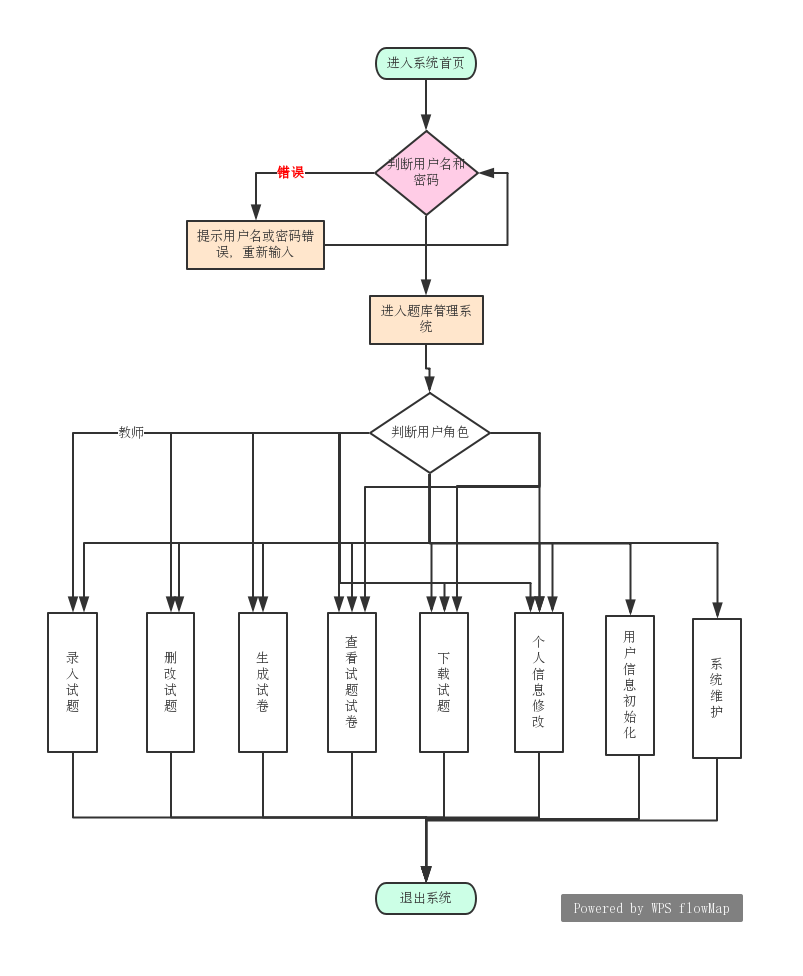
根据题库管理系统的需求分析，其业务流程图如图3.1所示****

图3.1

**3.1.2系统角色的划分及权限分配**

根据实际的教学需求可以将题库系统的用户分为三级系统管理员（0级）、教师（1级）、学生（2级），这三种级别的用户的权限由大到小，管理员（0级）拥有最高权限，学生（2级）权限最低。

1. 管理员：拥有该角色的账户在题库管理系统中具有最高权限。管理员负责题库管理系统的整体的运行与维护，可以执行题库管理系统提供的所有功能。包括教师以及学生初始信息的录入，题库的备份以及还原等。
2. 教师：教师的账号由管理员同一录入，后期可在个人中心进行修改。教师的主要职责是对自己录入的试题或者生成的试卷进行管理。
3. 学生：学生账号是由管理员同一录入，学生在题库管理系统中的权限最低，所能执行的功能也是最少的，只能查询试题，试卷的信息，或者下载试卷。

根据以上信息各角色权限及其可执行功能如下表所示：

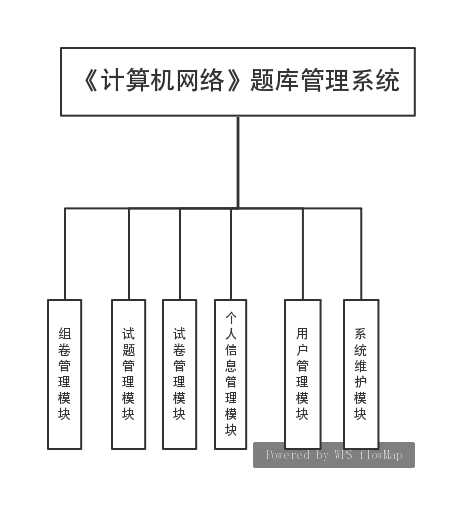
各角色及其可执行功能一览表3.1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 功能 | 功能名称 | 管理员 | 教师 | 学生 |
| 用户管理 | 教师/学生信息初始化 | 1 | 0 | 0 |
| 信息修改 | 1 | 1 | 1 |
| 登录 | 1 | 1 | 1 |
| 注销 | 1 | 1 | 1 |
| 试题管理 | 试题管理 | 1 | 1 | 0 |
| 试题录入 | 1 | 1 | 0 |
| 试卷管理 | 试卷管理 | 1 | 1 | 0 |
| 试卷下载 | 1 | 1 | 1 |
| 系统维护 | 试题/试卷备份还原 | 1 | 0 | 0 |
| 组卷管理模块 | 自动组卷 | 1 | 1 | 0 |
| 手动组卷 | 1 | 1 | 0 |
| 查询模块 | 用户信息查询 | 1 | 0 | 0 |
| 试题查询 | 1 | 1 | 1 |
| 试卷查询 | 1 | 1 | 1 |

**3.2题库系统对安全性的需求**

1. 题库系统本身的安全要求：为了保障系统的自身的安全在编写代码时采用安全性较高的代码，比如为了防止SQL注入在编写sql语句时采用特殊字符过滤以及传入参数。为了防止越权访问在每一个可执行功能前都要加上登录验证和访问权限的验证。
2. 题库系统数据的安全要求：为了保障题库系统中数据的安全，系统需提供备份还原功能，以保证在必要的时候可以对数据库中的所有数据进行备份以及还原。
3. **《计算机网络》题库管理系统功能模块设计及实现**

根据需求分析题库管理系统有如下的功能模块，如图4.1



**4.1《计算机网络》题库管理系统详细功能设计**

**4.1.1用户管理模块**

用户管理模块的主要功能是管理可以登录到题库管理系统账户，包括：

1. 教师/学生信息管理：初始化教师以及学生的账户，管理教师及学生的账户；
2. 信息修改：所有可以登录到题库管理系统中的用户都可以对自己的信息进行修改；
3. 登录：通过身份验证的用户可以登录到系统中使用系统提供的功能；
4. 注销：退出登录。

**4.1.2试题管理模块**

用于管理试题题库，在本系统中各类型试题分表存放，功能包括：

1. 试题管理：对试题进行增删改等操作；
2. 试题录入：手工录入试题，或者批量导入试题。

**4.1.3试卷管理模块**

试卷管理模块的功能主要是收录已经生成的试卷，为用户提供浏览和下载试卷的功能。

**4.1.4系统维护模块**

系统维护模块依据题库具体的使用情况以及教学安排等对试题题库和试卷进行备份和还原，为教学工作提供方便。

**4.1.5组卷管理模块**

组卷管理模块主要的功能是按照一定的需求生成试卷，该模块包含：

1. 手动组卷：教师可以手动录入试题也可以根据自己的需求从是题库中选择相应的试题进行组卷；
2. 自动组卷：教师根据自己的需求填写组卷规则，系统根据组卷规则生成相对应的试卷。

**4.1.6查询模块**

本模块主要是提供简单的查询功能，以便教师或学生等对试题试卷进行简单的查寻以及阅览。

**4.2《计算机网络》题库管理系统功能实现流程图**

**4.2.1用户管理模块**

**4.2.2试题管理模块**

用于管理试题题库，在本系统中各类型试题分表存放，功能包括：

1. 试题管理：对试题进行增删改等操作；
2. 试题录入：手工录入试题，或者批量导入试题。

**4.2.3试卷管理模块**

试卷管理模块的功能主要是收录已经生成的试卷，为用户提供浏览和下载试卷的功能。

**4.2.4系统维护模块**

系统维护模块依据题库具体的使用情况以及教学安排等对试题题库和试卷进行备份和还原，为教学工作提供方便。

**4.2.5组卷管理模块**

组卷管理模块主要的功能是按照一定的需求生成试卷，该模块包含：

1. 手动组卷：教师可以手动录入试题也可以根据自己的需求从是题库中选择相应的试题进行组卷；
2. 自动组卷：教师根据自己的需求填写组卷规则，系统根据组卷规则生成相对应的试卷。

**4.2.6查询模块**

本模块主要是提供简单的查询功能，以便教师或学生等对试题试卷进行简单的查寻以及阅览。

**4.3 数据库设计**

**4.3.1数据库概念结构设计**

概念模型是数据库概念结构设计中不可或缺的重要组成部分，为数据库提供基础性的阐释性的结构体系。概念模型的表示方法较多，最常用的模型就是E-R图模型，以下是题库管理系统中设计到的概念模型：（E-Rt图）

**4.3.2数据库物理结构设计**

数据库逻辑结构设计的实质就是将已有的E-R图转换成与之相适应的数据模型的过程。在这个阶段，就是要把概念模型经过映射、优化后获取完善的数据库表结构。如果实体间是一对一的联系，直接把联系映射成表的一部分；如果实体间是一对多的联系，把联系映射成表，将多端的主键作为表的键；如果实体间是多对多的联系，先分解成一对多的联系，再映射成表。

**4.3.3数据库逻辑结构设计（各表及其字段所带表的意思）**

1. **《计算机网络》题库管理系统组卷算法的的设计**

**5.1 遗传算法概述**

遗传算法是模拟达尔文的遗传选择和自然淘汰的生物进化过程的计算模型，是一种通过模拟自然进化过程搜索最优解的方法，它是用来解决多约束条件下的最优问题。

遗传算法是从代表问题可能潜在的解集的一个种群开始的，而一个种群则由经过基因编码的一定数目的个体组成。每个个体实际上是染色体带有特征的实体。染色体作为遗传物质的主要载体，即多个基因的集合，它决定了个体的形状的外部表现。因此，在一开始需要实现从表现型到基因型的映射即编码工作。由于仿照基因编码的工作很复杂，往往进行简化，如二进制编码，初始种群产生之后，按照适者生存和优胜劣汰的原理，逐代演化产生出越来越好的近似解，在每一代，根据问题域中个体的适应度大小挑选个体，并借助于自然遗传学的遗传算子进行组合交叉和变异，产生出代表新的解集的种群。这个过程将导致种群像自然进化一样的后生代种群比前代更加适应于环境，末代种群中的最优个体经过解码，可以作为问题近似最优解。

遗传算法提供了一种求解复杂系统优化问题的通用框架。它不依赖于问题的具体领域，对问题的种类有很强的鲁棒性，所以广泛应用于很多学科。遗传算法的主要应用领域有：函数优化、组合优化、生产调度问题、自动控制、机器人自动控制、图像处理和模式识别、人工生命、遗传程序设计、机器学习。

**5.1.1遗传算法的组成要素**

遗传算法有三个基本操作：选择、交叉、变异。这些操

作又有不同的方法来实现。

（1）选择。选择是用来确定重组或交叉的个体，以及

被选个体将产生多少子个体。首选计算适应度：按比例的适应度计算；基于排序的适应度计算。适应度计算之后是实际的选择，按照适应度进行父代个体的选择。可以挑选以下算法：轮盘赌选择、随机遍历抽样、局部选择、截断选择、锦标赛选择。

（2）交叉。基因重组是结合来自父代交配种群中的信息产生新的个体。依据个体编码表示方法的不同，可以有以下的算法：实值重组；离散重组；中间重组；线性重组；扩展线性重组。二进制交叉、单点交叉、多点交叉、均匀交叉、洗牌交叉、缩小代理交叉。

（3）变异。交叉之后子代经历的变异，实际上是子代基因按小概率扰动产生的变化。依据个体编码表示方法的不同，可以有以下的算法：实值变异、二进制变异。

**5.2基于遗传算法的组卷设计和实现**

**5.2.1组卷的编码方案**

用遗传算法求解问题， 首先要将问题的解空间映射成一组代码串，即染色体的编码问题。在传统的遗传算法中采用二进制编码。用二进制编码时，题库里的每一道题都要出现在这个二进制位串中，1 表示该题选中，0 表示该题未被选中。这样的二进制位串较长，且在进行交叉和变异遗传算子操作时，各种题型的题目数量不好控制。采用实数编码方案， 将一份试卷映射为一个染色体，组成该试卷的每道题的题号作为基因，基因的值直接用试题号表示，每种题型的题号放在一起，按题型分段，在随后的遗传算子操作时也按段进行，保证了每种题型的题目总数不变。比如，要组一份《计算机网络》试卷，其选择题 10 道，填空题 5 道，简答题 6道，则染色体编码是：

(10、11、34、66、7、170、76、23、52、101 | 52、36、67、11、123 | 99、85、45、56、44、34)

选择题 | 填空题 | 简答题

试卷初始种群不是采用完全随机的方法产生，而是根据题型比例、总分、知识点不重复等要求随机产生，使得初始种群一开始就满足了题型、总分、答题时间和知识点等要求。这样加快遗传算法的收敛并减少迭代次数，由于不同的题型是从不同的题型表中取出，有可能在同一个基因串中会出现相同的试题编号，因它们属于不同题型，故这种情况是正常的，不影响组卷。采用分组实数编码，可以克服以往采用二进制编码搜索空间过大和编码长度过长的缺点，同时取消了个体的解码时间，提高了求解速度。

**5.2.2适应度函数的计算方案**

适应度函数是用来评判试卷群体中个体的优劣程度的指标，遗传算法利用适应度值这一信息来指导搜索方向，而不需要适应度函数连续或可导以及其它辅助信息。因为时间、章节等要求在初始化种群时已经考虑，这里只剩下难度系数要考虑的了。所以适应度函数由试卷难度系数公式转换而成。试卷难度系数公式：P=∑Di×Si/∑Si；其中 i=1，2，....N，N 是试卷所含的题目数，Ｄi，Ｓi 分别是第 i 题的难度系数和分数。用户的期望难度系数 EP 与试卷难度系数 P 之差 f越小越好，f=|EP-P|， 而适应度函数是值越大越好，因此将目标函数 f 转换成适应度函数 F=e-0.03f，这种采用加权误差的适应度函数可以较好地反映求解组卷问题的特征，当试卷个体对各项组卷约束条件的误差越小时，它的适应度函数值就越大，则表示试卷个体越接近卷目标。

**5.2.3选择算子**

选择算子的作用在于根据个体的优劣程度决定它在下一代是被淘汰还是被复制。通过选择，将使适应度高的个体有较大的生存机会。本系统采用轮盘赌方法，它是目前遗传算法中最常用也是最经典的选择方法。其具体实现为：规模为 M 的群体 P={a1、a2、… am} ，其被选择概率为：

**5.2.4交叉算子**

由于在编码时采用的是分段实数编码，所以在进行交叉时采用分段单点交叉（按题型分段来进行交叉），整个染色体就表现为多点交叉。交叉的实现过程：将群体中的染色体任意进行两两配对，对每对染色体产生一个[0， 1 ]的随机数 r， 若 r≤pc（据经验，pc 值取 0.6 到 0.8），则分段随机产生一个交叉点，然后分段进行互换以得到下一代。

交叉后生成的子代有可能因存在重复的题号而非法。出现这种情况要将出现的题号换成该段中没有出现过的题号， 这样重新得到新子代。

**5.2.5突变算子**

在遗传算法中，变异概率一般较小。这里不分段进行变异，而是只对某段上的某个基因进行变异。对某个染色体，随机生成一个[0，1]范围内的实数 r，若r≤pm（据经验，pm 值取 0.01 到 0.02），则对该染色体进行变异，否则不进行变异。变异的操作如下：在[1，n]范围内随机生成一个段号 f，设该段的段长为 L，则在[1，L]范围内随机生成一个变异位置 P，以一定的原则从题库中选择一个变异基因，变异基因的选择原则为：与原基因题型相同的；满足答题时间相同；考试章节相同等要求。

**5.2.6停止条件**

1. **《计算机网络》题库管理系统测试与实现**

**6.1系统功能的实现效果**

**6.2系统测试**

**6.3错误处理**

1. **展望与总结**

**参考文献**

**致谢**

**附录**