

DOI:10.19651/j.cnki.emt.1802080

# 一种分阶求解回溯组卷算法\*

李川<sup>1</sup> 张少茹<sup>2</sup>

(1. 西安航空学院 计算机学院 西安 710077; 2. 西安交通大学 西安 710061)

**摘要:** 针对在线测评的广泛应用及现有组卷算法组卷效率低、过程繁琐、无法确保质量等问题,对人工组卷、随机函数选取法、回溯试探法及遗传算法进行了研究,并在回溯试探法的基础上提出一种分阶求解回溯组卷算法。首先依据分阶求解剪枝的思想,建立算法模型,再以提高效率和质量为目标,建立算法分析模型,按照题型、章节等指标将题库分阶求解,提高组卷效率及组卷质量。试验验证的结果表明,随着题库题目数量的增加,分阶求解回溯组卷算法在组卷效率、组卷质量(重复率、覆盖率)、试卷质量(区分度、信度)等方面都有显著提高。

**关键词:** 组卷;回溯;分阶;剪枝

**中图分类号:** TN98    **文献标识码:** A    **国家标准学科分类代码:** 510.4030

## Hierarchical solving paper-generating algorithm based on backtracking search

Li Chuan<sup>1</sup> Zhang Shaoru<sup>2</sup>(1. College of Computing, Xi'an Aeronautical University, Xi'an 710077, China;  
2. Xi'an Jiaotong University, Xi'an 710061, China)

**Abstract:** In view of the problems that the online judge is widely used and the existing paper-generating algorithm is low in efficiency, process complexity, and the quality of the paper-generating is not high. Study the artificial paper-generating, random function selection method, backtracking search method and genetic algorithm, and proposes a Hierarchical Solving paper-generating algorithm based on backtracking search. First, establish an algorithm model according to the idea of hierarchical solving pruning, and then the algorithm analysis model is set up with the goal is to improve efficiency and quality of the paper-generating, according to the type of questions, chapters and other indicators, the question bank is hierarchical solved to improve the efficiency of the paper-generating and the quality of the paper-generating. The experimental results show that with the increase of the amount of data, the hierarchical solving paper-generating algorithm based on backtracking search has obviously improved in the efficiency of paper-generating(repetition rate, coverage rate), the quality of paper-generating and the quality of paper(discrimination, reliability).

**Keywords:** paper-generating; backtracking; hierarchical; pruning

## 0 引言

从古到今,考试都是衡量知识能力水平的最有效形式<sup>[1]</sup>,随着教育信息化的发展,传统的教师命题纸质试卷的考试模式暴露出了很多弊端,例如浪费资源、过程繁琐、效率低、试题保密性不强、缺乏规范性和合理性等问题<sup>[2]</sup>。在全球互联网、信息技术的飞速发展,网络在线考试利用计算机网络的开放共享性,简化了考试流程,很多工作可是使用计算机自动完成,极大的减轻教师的工作,而且也节约

资源,同时提高了考试质量。

网络在线考试系统一般包括试题管理模块、考试管理模块、组卷模块、考试管理模块、成绩管理模块等<sup>[3]</sup>,其中网络在线考试系统的核心和技术难点是组卷模块,组卷就是按照考试要求,由计算机根据某种组卷策略自动从现有试题库中抽取试题,组成若干套试卷,组成的试卷要科学、合理且试题无规可寻。组卷的关键是组卷策略(算法)的设计,衡量组卷策略的依据是组成试卷的质量(科学合理性)、重复率及组卷效率等。

收稿日期:2018-09-13

\* 基金项目:国家自然科学基金(71373203)、陕西省科技工业攻关项目(2016GY-139)、西安航空学院校级科研基金项目(2016JS3223)资助

组卷策略很多,常用的组卷策略有人工组卷、随机函数选取法、回溯试探法、遗传算法、蚁群算法等<sup>[4-6]</sup>。

1)人工组卷

由教师根据考试规则(总分、题目类型、数量、难度等),从题库中手工抽取合适的试题来组卷。显然,人工组卷方式主观性太强,只能当做纸质试卷命题的一种简单替换,当题库较大或组卷份数较多时,人工组卷的质量和效率都很低。

2)随机函数选取法

根据组卷要求由计算机采用随机函数反复的在题库中选择一个符合要求并且没有重复的试题组卷,直到组卷成功或因为题库试题有限等原因无法满足组卷要求导致组卷过程停止。随机函数选取法策略简单,在题库规模小、组卷份数少的情况下组卷效率较高,但该方法组卷质量难以保证,在题库规模较大时无法满足组成试卷的科学合理性,且组卷消耗效率低、占用资源大<sup>[7]</sup>。

3)回溯试探法

回溯法适用于求一个问题的解集(全部解)或符合某些约束的最优解的情况<sup>[8]</sup>。

回溯试探法采用随机函数抽取法选题,在选题过程中记录选题步骤,当验证所选题不能满足既定目标条件且组卷过程未结束时,则根据记录的选题步骤过程回溯,即终止本条路径退回一步重新抽题,通过不断地回溯试探抽题直到组卷成功。回溯试探法在组卷过程中终止部分搜索,避免了无效搜索,较随机函数选取法在性能有了明显提高。因为回溯试探过程中需要记录每一步的状态,当题库规模较大时,需要记录的状态将非常巨大,即占用空间较多,而反复的回溯也会增加时间复杂度,因此该算法组卷成功率高,但性能不高<sup>[9]</sup>。

4)遗传算法<sup>[10-11]</sup>

遗传算法模拟生物界的进化过程,先生成初始群体,通过演变,完成个体重组,再计算初始群体中个体的适应度,选取适应度高的优良个体组成新一代,如此反复迭代,从而找到满足组卷要求的最优解<sup>[12-13]</sup>。这种算法组卷成功率高,但算法实现困难。

上述算法对于设计组卷策略各有优缺点,本文针对题库规模大、约束条件多的情况研究设计一种分阶求解回溯组卷算法模型,以解决常见算法组卷质量低、效率低、实现困难等问题。

1 分阶求解回溯组卷算法设计

回溯试探算法(探索与回溯)是一种决策树(decision tree method,DFT)选优搜索算法,为了达到目标,根据约束条件向前搜索,并且保存记录每次搜索得状态以保证回溯到指定状态,当搜索到某一状态时,发现继续搜索将达不到目标时,就回退到上一状态继续搜索,直到达到目标或遍历结束<sup>[14]</sup>。达到目标的状态集合就是问题的解,算法过程描述如图 1 所示。

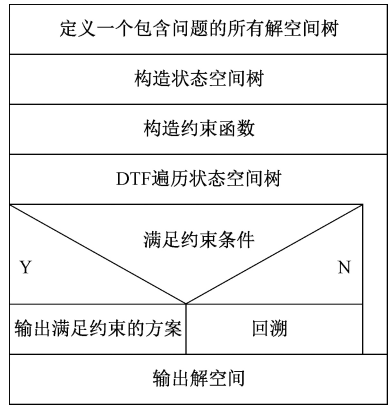


图 1 回溯试探过程 NS 图

从图 1 可以看出,回溯试探法在问题的状态空间树中,按 DTF 从根结点开始搜索状态空间树,每次搜索判断结点是否满足约束条件。如果不满足则回溯,且跳过该节点的子树的遍历,称为“剪枝”<sup>[15]</sup>;否则,以 DTF 策略遍历该节点的子树。在该算法中,影响空间复杂度的主要因素是遍历的状态点,影响时间复杂度的主要因素是空间状态树遍历次数,要记录的状态数据量很小,因此空间复杂度影响较小,而对于时间复杂度而言,如果能尽早的剪掉无解的子树,即可避免大量无效的遍历,本文对此提出了一种分阶求解回溯组卷算法,算法流程如图 2 所示。

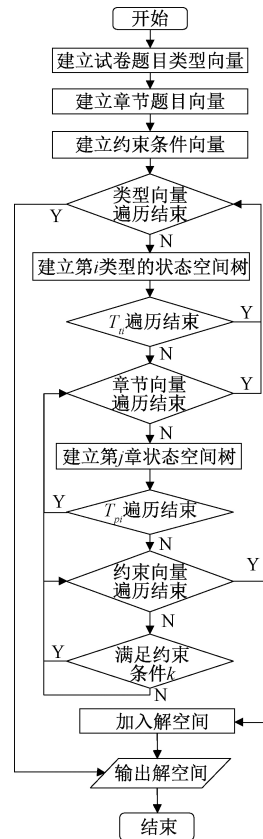


图 2 分阶回溯组卷流程

一个复杂问题的解空间一般是由多个简单问题的解空间构成。这些简单问题解空间构成一个决策序列。该复杂问题的所有可能的决策序列就是这个问题的解空间。但解空间中并非所有的解都是目标解,可行解可能是目标解,前提是满足约束。回溯就是在解空间中找到可行解的过程,达到目标的解也可能有多个,即有多个可行解,对于组卷而言,一般还需要找到最优解,最优解是指在约束条件下使目标值达到最优的可行解。一般试卷都会分为多种类型的题目,可以按照题目类型对每种类型建立状态空间树,并将约束条件分解到每种类型的题目上,再进行遍历判断,对于不满足约束条件的结点进行剪枝。采用同样的过程再按章节对空间状态树进行剪枝。这样可以尽早的去掉无效的遍历,提高搜索效率。

具体算法过程如下:

### 1) 初始化

(1) 读取  $n$  个类型的试题向量  $\vec{T}(t_1, t_2, t_3, t_4, \dots, t_i, t_{i+1}, \dots, t_{2n-1}, t_{2n})$  其中,  $t_i (\{i | i = 2k+1, k \in \mathbf{Z}\})$  表示第  $i$  类试题的数量,  $t_{i+1}$  表示第  $i$  类试题分值,且满足  $S_p = \sum_{i=1}^n t_i \times t_{i+1}$ ,  $S_p$  表示试卷总分。

(2) 读取  $m$  个章节的试题向量  $\vec{P}(p_1, p_2, p_3, p_4, \dots, p_j, p_{j+1}, \dots, p_{2m-1}, p_{2m})$ , 其中,  $p_j (\{j | j = 2k+1, k \in \mathbf{Z}\})$  表示第  $j$  章节试题的数量,  $p_{j+1}$  表示第  $j$  章节试题占比,且满足  $p_{j+1} = \frac{p_j}{\sum_{j=1}^m p_j}$ 。

(3) 读取  $l$  个约束条件向量  $\vec{C}(c_1, c_2, \dots, c_k, \dots, c_l)$ , 其中,  $c_k$  表示第  $k$  个约束条件。组卷成功状态与约束条件  $c_k$  之间存在某种函数关系,记为  $f(c_k) = \begin{cases} 1 & \text{组卷成功} \\ 0 & \text{组卷失败} \end{cases}$ 。

### 2) 遍历类型向量

遍历  $\vec{T}(t_1, t_2, t_3, t_4, \dots, t_i, t_{i+1}, \dots, t_{2n-1}, t_{2n})$ , 生成  $n$  个空间状态树集合  $\mathbf{T}_i \{t_i, i \in N\}$ ,  $t_s$  表示第  $i$  种类型试题的空间状态树。

### 3) 遍历空间状态树 $t_{ii}$

当状态树有解时,遍历章节向量  $\vec{P}(p_1, p_2, p_3, p_4, \dots, p_j, p_{j+1}, \dots, p_{2m-1}, p_{2m})$ , 生成  $m$  个空间状态树集合  $T_p \{t_p, j \in N\}$ ,  $t_p$  表示第  $j$  章节试题的空间状态树。

### 4) 遍历空间状态树 $t_p$

当状态树有解时,遍历约束条件向量  $\vec{C}(c_1, c_2, \dots, c_k, \dots, c_l)$ , 判断如果函数  $f(c_k) = 1$ , 满足继续遍历约束条件向量  $\vec{C}(c_1, c_2, \dots, c_k, \dots, c_l)$ , 否则不满足转向步骤 3)。

### 5) $k \geq l$

如果满足  $k \geq l$ , 将解写入集合  $C_j \{c_{j_h}, h \in N\}$  中, 其中  $c_{j_h}$  表示第  $j$  章第  $h$  组满足约束的试题组合。

### 6) $j \geq m$

如果满足  $j \geq m$ , 将解集合  $C_j$  写入解集合  $C_i \{c_{i_j},$

$j \in N\}$  中, 其中  $c_{i_j}$  表示第  $i$  种题型第  $j$  章解集合。

### 7) $i \geq n$

如果满足  $i \geq n$ , 循环嵌套遍历集合  $C_i$ , 输出满足条件的解集合。

利用分阶求解回溯组卷算法也可以在根据试题难度或其他指标再构造子空间状态树进行求解, 但一般不要超过三级分阶, 否则可能会导致算法效率下降的情况。

## 2 算法分析及测试

### 2.1 算法分析模型

衡量一个组卷算法的优劣既要看算法的效率, 还要看组卷的质量。可根据算法的平均执行时长、试题重复率、试卷覆盖率、试题区分度、信度等指标建立组卷算法分析模型。

定义一次组卷成功执行时长  $L$  为:

$$L = T_e - T_s$$

式中:  $T_e$  表示组卷完成时间;  $T_s$  表示组卷开始时间。

定义平均组卷时长  $\bar{L}$  如式(1)。

$$\bar{L} = \frac{\sum_{i=1}^n L_i}{n} \quad (1)$$

定义试题重复率。设  $N_p$  表示每套试卷题目数,  $N$  表示总试卷数,  $f_i$  表示题库中试题  $i$  在所有试卷中出现的频率,  $f_i$  计算如式(2)所示。

$$f_i = \frac{\sum_1^N \sum_1^{N_p} f(i)}{N} \quad (2)$$

其中  $f(i)$  计算如式(3)。

$$f(i) = \begin{cases} 1 & \text{题库中第 } i \text{ 题与某试卷中一题相同} \\ 0 & \text{题库中第 } i \text{ 题与某试卷中一题不相同} \end{cases} \quad (3)$$

试题重复率  $R_i$  是指每套试卷中任一试题在所有试卷试题中出现的比率, 计算如式(4)。

$$R_i = \frac{\sum_1^{N_p} f_i / N}{N_p} \quad (4)$$

试卷覆盖率  $R_c$  是指试卷试题所属指标占题库试题总指标的比率, 计算如式(5)。

$$R_c = \frac{C_p}{C_d} \quad (5)$$

式中:  $C_p$  表示试卷中所属指标数量;  $C_d$  表示题库所有指标数量。

平均覆盖率  $\bar{R}_c$  计算如式(6)。

$$\bar{R}_c = \frac{\sum_1^N R_c}{N} \quad (6)$$

覆盖率标准偏差  $S$  计算如式(7)。

$$S = \sqrt{\frac{\sum_{i=1}^N (R_{c_i} - \bar{R}_c)^2}{N - 1}} \quad (7)$$

试题/试卷区分度表示试题/试卷对考生真实能力水平的衡量<sup>[16]</sup>。

某个试题的区分度的计算过程如下。

将所有第  $i$  题考分降序排序,统计前 27% 部分的平均分  $H_i$ ,统计后 27% 部分的平均分  $L_i$ ,设第  $i$  题满分为  $W_i$ ,第  $i$  题区分度  $d_i$  计算如式(8)。

$$d_i = \frac{H_i - L_i}{W_i} \quad (8)$$

一套试卷的区分度的计算过程如下。

将所有考分降序排序,统计前 27% 部分的平均分为  $H$ ,统计后 27% 部分平均分为  $L$ ,试卷满分为  $W$ ,试卷区分度  $d$  计算如式(9)。

$$d = \frac{H - L}{W} \quad (9)$$

所有试卷的平均区分度  $\bar{d}$  计算如式(10)。

$$\bar{d} = \frac{\sum_{i=1}^N d}{N} \quad (10)$$

区分度标准偏差  $S_d$  计算如式(11)。

$$S_d = \sqrt{\frac{\sum_{i=1}^N (d - \bar{d})^2}{N - 1}} \quad (11)$$

信度指考试结果的可靠性。信度  $\alpha$  计算如式(12)。

$$\alpha = \frac{N_p}{N_p - 1} \left[ 1 - \frac{\sum_{i=1}^{N_p} S_i^2}{s^2} \right] \quad (12)$$

式中: $N_p$  表示一套试卷的题目总数; $S_i^2$  表示所有考生在第  $i$  题的方差; $s^2$  表示所有考生所得总分的方差。

2.2 试验分析

1) 算法效率、组卷质量分析

将分阶求解回溯组卷算法、传统回溯试探组卷算法、随机函数选取法及遗传算法分别应用于在线考试系统中,首先使用《计算机基础》题库,该题库包括 4 个类型题目:选择题、多项选择、填空题、判断题,每种类型 1 500 道题目,共 6 000 题。组卷要求如表 1 所示,试验结果如表 2 所示。

表 1 计算机基础组卷要求

序号	要求
1	满分 100 分
2	选择 40 道,每题 1 分;多选题 20 个,每题 2 分;填空题 10 道,每题 1 分;判断题 10 道,每题 1 分
3	平均难度 0.5
4	试题分布:第 1 章占 6%;2~3 章各占 12%;4~7 章各占 15%;第 8 章占 10%
5	概念题占 20%,综合应用题占 60%,计算题占 20%
6	试卷份数 100 套

通过试验结果可以看出,分阶求解回溯组卷算法在平

表 2 4 种算法在《计算机基础》题库中的试验比较

指标	传统回溯 试探组卷 算法	分阶求解 回溯组卷 算法	随机 函数 选取法	遗传 算法
$\bar{L}$	1.11	0.74	1.52	0.81
$R_r$	3.16%	2.85%	4.27%	2.93%
$\bar{R}_c$ (章节)	95%	100%	91%	100%
S(章节)	0.029	0	0.033	0
$\bar{R}_c$ (难度)	99.2%	98%	94%	99.5%
S(难度)	0.021	0.012	0.032	0.017
$\bar{R}_c$ (知识类型)	99.6%	98.3%	93.5%	97.6%
S(难度)	0.018	0.011	0.022	0.015

均组卷时间、重复率都有了明显提高,在平均覆盖率方面略有降低,但覆盖率标准偏差上明显提高,说明算法组卷结果比较稳定,随机函数选取法各方面表现最差,虽然遗传算法的各指标表现也良好,但其缺乏一定的稳定性。

采用相似的组卷要求在《高等数学》试题库中进行试验,该题库选择题、多项选择、填空题、判断题各 5 000 道题目,共 20 000 题,试验结果如表 3 所示。

表 3 4 种算法在《高等数学》题库中的试验比较

指标	传统回溯 试探组卷 算法	分阶求解 回溯组卷 算法	随机 函数 选取法	遗传 算法
$\bar{L}$	1.89	0.83	2.46	0.86
$R_r$	1.47%	1.25%	2.33%	1.23%
$\bar{R}_c$ (章节)	97.3%	100%	95%	100%
S(章节)	0.016	0	0.021	0
$\bar{R}_c$ (难度)	99.9%	99.5%	94%	100%
S(难度)	0.012	0.009	0.017	0
$\bar{R}_c$ (知识类型)	99.9%	99.6%	94.9%	99.7%
S(难度)	0.011	0.009	0.018	0.01

通过试验分析,当题库规模较大时,4 种组卷算法的效率下降,组卷质量提高,但分阶求解回溯组卷算法效率基本不变,组卷质量提高的幅度更大,遗传算法的测试效果与分阶求解回溯组卷算法的测试效果相当,但遗传算法的过程复杂,实现算法的难度远高于分阶求解回溯组卷算法。

2) 试卷质量分析

经试验结果表明,传统回溯试探组卷算法和随机函数选取法在组卷效率、组卷质量等方面都明显略于遗传算法和分阶求解回溯组卷算法,因此在试卷质量方面再对遗传算法和分阶求解回溯组卷算法进行比较分析。分别采用遗传算法和分阶求解回溯组卷算法在《高等数学》试题库中组成 4 套试卷对 4 个专业学生进行统考,其中,参加采用遗传算法组卷的考试人数 889 人,参加采用分阶求解回溯组卷算法的考

试人数 921 人,人数差距不大,考试结果如表 4 所示。

表 4 分阶求解回溯组卷考试结果

指标	遗传算法结果	分阶求解回溯组卷算法结果	参考值
$\bar{d}$	0.31	0.37	优( $>0.4$ ); 良( $[0.3,0.4]$ ); 可( $[0.2,0.3]$ ); 差( $<0.2$ )
$S_d$	0.16	0.14	
$\alpha$	0.702	0.753	$\alpha > 0.7$ 信度较高

通过考试结果可以看出,分阶求解回溯组卷算法试卷质量区分度属于良,信度属于优,区分度标准偏差较小,且各方面都要优于遗传算法的试卷质量。

### 3 结 论

组卷算法是在线考试系统中研究的重点问题,本文在研究了人工选取法、随机函数选取法、回溯试探法及遗传算法等常见组卷算法的各自优缺点后,针对回溯试探法组卷在题库规模较大时组卷效率较低的问题,依据早剪枝、去无效的思想,提出了分阶求解回溯组卷算法,按照题型、章节等指标将题库分阶求解,提高组卷效率及组卷质量。算法首先建立题目类型向量、章节向量、约束条件向量,再依次对 3 个向量集进行遍历,遍历时对遍历的每个向量建立状态空间树,如果满足所有约束,则是一个有效解,否则,终止本次遍历,跳过后续的遍历过程,即提前剪枝,提高了算法执行效率。

为了验证算法的效果,根据算法的平均执行时长、试题重复率、试卷覆盖率、试题区分度、信度等指标建立组卷算法分析模型。通过具体的试验测试,结果表明,随着题库题目数量的增加,常见的 4 种组卷算法的效率下降明显,组卷质量有所提高,但分阶求解回溯组卷算法效率基本不变,组卷质量提高的幅度更大,虽然遗传算法的试验效果与分阶求解回溯组卷算法的试验效果相当,但遗传算法的过程复杂,实现算法的难度远高于分阶求解回溯组卷算法。由此得出,分阶求解回溯组卷算法适用于规模较大的题库组卷。本算法的不足之处在于实现算法的过程中建立了 3 个向量集,当题库规模较大时,会占用一定的存储空间。

### 参考文献

- [1] 郑旭东,高守林,任友群.计算机化自适应测验及应用于规模化考试的主要问题[J].开放教育研究,2016,22(4):40-49.
- [2] 张华华.计算机化考试与中国教育评估[J].心理学探新,2013,33(5):387-391.
- [3] LIU L, SHI H, ZHAI R. The design and implementation of online exam system[J]. Applied Mechanics and

Materials,2014,3634(687):2506-2509.

- [4] OGAGA IGHOSE B, ADELEKE I A, DAMOS M, et al. Optimization of biodiesel production from Thevetia peruviana seed oil by adaptive neuro-fuzzy inference system coupled with genetic algorithm and response surface methodology [J]. Energy Conversion and Management, 2017, 132:231-240.
- [5] YAN L, SHUHONG L, XIURONG L. Test paper generating method based on genetic algorithm [J]. AASRI Procedia, 2012, 1(Complete):549-553.
- [6] 鲁萍,王玉英.多约束分级寻优结合预测计算的智能组卷策略[J].计算机应用,2013,33(2):342-345.
- [7] 刘晓东,蒋睿.分数阶 Kalman 滤波算法回溯长度选择方法研究[J].电子测量与仪器学报,2015,29(1):132-138.
- [8] BAI G, TIAN Z, ZUO M J. An improved algorithm for finding all minimal paths in a network [J]. Reliability Engineering & System Safety, 2016, 150:1-10.
- [9] 沈跃,丁灵卫,吴翊轩,等.回溯自适应匹配追踪电能质量信号重构方法[J].电子测量与仪器学报,2017,31(5):731-738.
- [10] 杨素锦,陈莹.基于改进遗传算法的自动组卷问题研究[J].煤炭技术,2012,31(12):217-219.
- [11] 刘怀兰,牛辉,王佳.基于改进遗传算法的智能组卷模型优化[J].华中科技大学学报(自然科学版),2013,41(5):82-85.
- [12] 曾学强,刘志刚,符伟杰,等.改进差分进化算法在电力系统无功优化中的应用[J].电网技术,2012,36(2):121-125.
- [13] 陈青青.递归随机分割算法在考试系统中的应用[J].计算机工程,2011,37(14):268-270,273.
- [14] 黄越岭,朱德全.复杂性与情境性的耦合:网络学习评价的新取向及其路径探求[J].电化教育研究,2013,34(7):47-52.
- [15] YI C, DU X Y, FENG S, et al. The design of course-oriented online judge [J]. Advanced Materials Research, 2012, 433-440:4736-4740.
- [16] 江文奇.基于方案区分度的组合权重信度判别及修正研究[J].系统工程与电子技术,2012,34(10):2090-2093.

### 作者简介

李川,硕士、讲师,主要研究方向为软件算法理论、大数据分析。

E-mail:lch9964121@163.com

张少茹,博士、教授、博士生导师,主要研究方向为慢性病、传染性疾病防治管理。

E-mail:shaoru@126.com