全国第三届研究生数学建模竞赛

题 目　　　学生面试问题

北京理工大学：李然，王继辉，王建秋

摘　　　要：

本文通过使用图论、局部搜索算法确定了给定条件下的老师数量的最小值 M，建立了分配老师的组合优化模型，并求解模型给出具体分组情况，最后提出了改进的公平性指标。

问题 1：论文通过解决对偶命题来实现的，使用图论的边、度关系知识，获取 M 的一个较优下界(M=$\left\lceil \frac{1+\sqrt{1+48n}}{2} \right\rceil$，同时回答了 M mod 3 条件下 M 的更精确的下界)，然后利用这个下界，使用局部搜索算法，搜索完全图 Km 中的无边重复的 K_4 的个数，从而找到最小的 M。

问题 2：将问题抽象为一个 TTP（Time Table Planning)，建立了分配老师与学生的组合优化模型，通过设定约束条件的优先级别，确定了优化目标，并利用遗传算法对模型进行求解，给定 M,N 的“面试组”方案（结果详见正文）。

问题 3：前半部分的思路与问题 1 相同，M 的一个较优下界为不小于 $\left\lceil \frac{1+\sqrt{1+48n}}{2} \right\rceil$ 的偶数，同时在局部搜索算法中多引入一个状态矩阵 S，而且还研究了此条件下文（理）老师分配“面试组”时的最大利用率近似为 $\frac{2}{3}$；后半部分的思路与问题 2 相同，修改问题 2 已建立的模型及优化目标。

问题 4：从面试的均匀性、公平性和信度出发，考虑到当前学科交叉和学生的综合素质日益重要，我们提出新的分配方案--每 4 位老师同时面试 3 位学生，每位学生随机参加 3 场面试。

关键字：对偶　完全图　较优下界　局部搜索算法　TTP　遗传算法　公平性指标

问题 1

设 G 为 m 阶无向简单图，若 G 中的每个定点均与其余的（m-1）个顶点相邻，则称 G 为 m 阶无向完全图，记为 Km. 如四阶完全图 K4。

如果用 G 的每个顶点来表示不同的老师，用 G 中的边来表示老师在同一个“面试组”这一关系，则 G 中无边重复的 K4 图，就对应了一个“面试组”方案，同时，每有一个面试方案，就意味着老师可以接受一个考生的面试请求。于是本问题就等价于下面一个对偶的图论问题：

对偶命题 1：设 G 是一个 m 阶无向完全图 Km，n 为 G 中无重复边的 K4 的个数，则当 n= N 时，求能够满足条件的最小的 m 值。

为解决的这个问题，可以通过下面两个步骤：

一、 确定 m 的一个较优的下界

引理 1:对于一个 m 阶完全图 G，n 表示 G 中无重复边的 K4 的个数，则必有：

$$m \geq \frac{1+\sqrt{1+48n}}{2}$$

证明：对于 G 来说，从中每删除一个 K4 子图，G 的边就将减少 6 条。则 G 能提供的 K4 的最大个数为：$\frac{1}{6}C_m^2$,即$\frac{1}{12}m(m-1)$，

进一步有 $\frac{1}{6}C_m^2 \geq n$

化简后 $m^2 - m - 12n \geq 0$

则有 $m \geq \frac{1+\sqrt{1+48n}}{2}$

上面虽然考虑了 K4 与 Km 的 6 条边的对应关系，但没有分离出不能够组成一个 K4 的情况，则得到的 m，比要求的 M 要小，对于 m，n 的关系描述比较粗略。为了更细致的刻画这一点，下面基于 m 的不同情形进行讨论：

引理 2：对于一个 m 阶完全图 G，n 表示 G 中无重复边的 K4 的个数,则有：

（1）m mod 3 = 0 时，$m \geq \frac{3+\sqrt{9+48n}}{2}$

（2）m mod 3 = 2 时，$m \geq \frac{2+\sqrt{4+48n}}{2}$

（3）m mod 3 = 1 时，$m \ge \frac{1+\sqrt{1+48n}}{2}$

证明：对于 G 中任意顶点 v，有 v 的度为 d（v）= m-1；

（1） 由于 m mod 3 = 0，则 G 删除若干个无边重复的 K4 图后，至少要剩余 m 条边，无法被使用。则有

$$\frac{1}{6}(C_m^2 - m) \ge n$$

化简后得：$m^2 - 3m - 12n \ge 0$

则有：$m \ge \frac{3+\sqrt{9+48n}}{2}$

（2） 由于 m mod 3 = 2，则从 G 删除若干个无边重复的 K4 图后，至少要剩余 m/2 条边，无法被使用。则有

$$\frac{1}{6}(C_m^2 - m/2) \ge n$$

化简后：$m^2 - 2m - 12n \ge 0$

则有：$m \ge \frac{2+\sqrt{4+48n}}{2}$

（3） 与引理 1 相同，证明略

现在比较一下，通过前面的引理求出的 m 与实际的 m.可以直观的看出两者之间的差距。

A：实际关系对照表

M	4	5	6	7	8	9	10	11
N	1	1	1	2	2	3	5	6

B：较优 M/N 关系反映表

M（0）	6/5.27	9/6.62	9/7.68	12/8.58	12/9.38	12/10.11	12/10.78	15/11.41
M（2）	5/4.60	8/6	8/7.08	11/8	11/8.81	11/9.54	11/10.21	14/10.84
M（1）	4/4	7/5.42	7/6.52	10/7.44	10/8.26	10/9	10/9.67	13/10.31
N	1	2	3	4	5	6	7	8

（上表单元格“/”左边为对应的最近不小于右边的整数，右边为原始值）

从上面的对照中，我们可以看到，利用上面的 m，n 关系求得的结果与实际情况还有一定的误差。为了改进结果，使用如下搜索算法。

二、搜索最理想的 M 值

1 前提条件

为了使用下面的搜索算法，首先要将 *Km* 图用 0，1 数字矩阵表示。设 G 的邻接矩阵为 A，当 A[i][j] = 1，表示存在 i，j 两个顶点间存在一条直接

边；A[i][j] = 0，表示 i，j 两个顶点间没有直接的连接边；同时，A[i][i] = 0，表示 i 顶点与自身永远没有直接边相连。此时 Km 可以表示为：

$$Km = \begin{array}{c c} \begin{array}{c} 1 \\ 2 \\ 3 \\ \cdots \\ m \end{array} & \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix} \\ & \begin{array}{ccccc} 1 & 2 & 3 & \ldots & m \end{array} \end{array}$$

而 K4 可以表示为：

$$K_4 = \begin{array}{c c} \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \\ & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \end{array}$$

通过上面的结论，在给定 n 的条件下，可以获取对应的 m 的一个下界，从此时的 m 开始，不妨记为 M = m。

2 K_4 搜索算法：

判断 Km 中无边重复的 K_4 数目的方法是循环搜索，具体是：

按照先从左到右，再从上到下的顺序，从 Km 中找到第一个不为 0 的元素。不妨记其为 A[i][j]，这就选定了 K_4 的一条边，然后再从与位于 A[i[j]后面，并且位于同行的元素 A[i][s]，A[i][t],(j<s,t<=m 并且 $s \neq j$).现在已经找到 K4 的四个顶点了，然后这个新生成的矩阵与 K_4 是否可以匹配，即计算 A[j][s]，A[j][t],A[s][t]是否都为 1,如果出现一个为 0,则再找一组 A[i][s],A[i][t],继续上面的检测。如果，已经检测了 i 行的所有 A[i][s]，A[i][t]组合了，仍然没有找到一个 K4，则说明当前顶点已经不再被使用了，此时将删除第 i 行与第 i 列；如果找到一组 A[i][s]，A[i][t]，使得这个矩阵与 K_4 是匹配的，这时，记录当前的 i，j，s，t 值，同时修改邻接矩阵 A，使得这个新矩阵的元素全为 0。然后再找下一个为 1 的元素。同时记录考察点的新位置，检查过程同上。最后，可以将 Km，变成一个全 0 的矩阵。这样搜索结束，针对当前 M 值的所有 $K4$ 图都已经出现了。如果这时 $K4$ 的个数小于 n,说明当前的 M 太小了。将 M = M+1,然后再重复上面的过程，就可以计算出最优的 M 值了。

下面以 $K7$ 为例，来说明 $K4$ 搜索算法的过程：

刚开始 G 的邻接矩阵是：

$$A=\begin{array}{c}1\\2\\3\\4\\5\\6\\7\end{array}\begin{pmatrix}0&1&1&1&1&1&1\\1&0&1&1&1&1&1\\1&1&0&1&1&1&1\\1&1&1&0&1&1&1\\1&1&1&1&0&1&1\\1&1&1&1&1&0&1\\1&1&1&1&1&1&0\end{pmatrix}$$

1　2　3　4　5　6　7

第一个不为 0 的元素为 A[1][2]，下面再从 A[1][3]，A[1][4]，…，A[1][7] 中任选两个。这里假设选择了 A[1][3]，A[1][4]。从而可以得到一个由 1，2，3，4 组成的一个矩阵：

$$\begin{pmatrix}0&1&1&1\\1&0&1&1\\1&1&0&1\\1&1&1&0\end{pmatrix}$$

这里恰好是一个 $K4$ 图，此时将上矩阵中为 1 的地方，都改成 0，现在 G 变成

$$A=\begin{array}{c}1\\2\\3\\4\\5\\6\\7\end{array}\begin{pmatrix}0&0&0&0&1&1&1\\0&0&0&0&1&1&1\\0&0&0&0&1&1&1\\0&0&0&0&1&1&1\\1&1&1&1&0&1&1\\1&1&1&1&1&0&1\\1&1&1&1&1&1&0\end{pmatrix}$$

1　2　3　4　5　6　7

此时第一个不为 0 的元素为 A[1][5]，后面只有 A[1][6]，A[1][7] 了。这时又可以得到一个

$$\begin{pmatrix}0&1&1&1\\1&0&1&1\\1&1&0&1\\1&1&1&0\end{pmatrix}$$

这里恰好也是一个 $K4$ 图，此时将上矩阵中为 1 的地方，A 中都改成 0，现在 G 变成

$$A = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ \\ \end{array}\begin{array}{c}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix} \\ \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{matrix}\end{array}$$

继续向下搜索，发现 A[2][5]=1，不为 0，以其为考察点，后面只有 A[2][6]，A[2][7]，由 2，5，6，7 组成一个矩阵

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

不能满足 $K4$，继续考察 A[2][5]的其他组合方案，但此时已经没有可以再选择的了。也就是说，现在即便 G 中还存在 $K4$，A[2][5]也不可能出现了，所以，将 A 矩阵中对应的 1 变成 0。

G 又变为：

$$A = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ \\ \end{array}\begin{array}{c}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{pmatrix} \\ \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{matrix}\end{array}$$

不断循环下去，…,直到 G 变成一个 7 阶零图。

$$A= \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ \\ \end{array}\begin{array}{c}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{matrix}\end{array}$$

A 变成一个 n 阶零矩阵,这时就找到 Km 中所有的无边重复的 K4 图。

搜索结束后，当 M=7 时，最多可以构造两个“面试组"，如果想获得 3 个，甚至更多“面试组”时,只需 M = 8，9，…，不断加 1 即可.

问题2分析

问题的优化目标是求得相对较优解，从而尽可能保证在面试过程中的公正性。此问题属于典型的组合优化问题，在组合过程中增加约束条件，分析约束条件Y1，Y2是保证面试老师之间工作量的均衡性，从而减少因为人为主观因素对面试结果的影响，这两个条件对最后公平的影响因子较大，因此在建模过程中可以考虑将其设为硬约束，而对于Y3，Y4条件，当面试学生数远远大于面试老师总数的时候，“面试组”成员之间不可避免的会出现一定量的重复，因此考虑将Y3，Y4定为模型的软约束，在保证硬约束的前提下尽量满足软约束的要求。
通过分析由于每个“面试组”成员数固定，而且每位考生有且只有一个“面试组”对其面试。
如图示：（T：Teacher；S：Student）

S T	S_1	S_2	S_3	S_4	S_5	……	S_N
	T_1	T_2	T_{10}	T_6	T_5	……	T_i
	T_2	T_5	T_8	T_{17}	T_{19}	……	T_j
	T_5	T_6	T_3	T_{19}	T_3	……	T_k
	T_7	T_{10}	T_{14}	T_{20}	T_{13}	……	T_l

问题求解就是将M位老师分别填入一张4×N的二维表中，由此可见此问题可以抽象出一个二维时间表问题（TTP：Time Table Problem）。

70年代中期，S.Eveo等人论证了课表问题是NP完全问题，对于NP完全问题至今没有找到线形时间可以解决的算法，采用穷举搜索算法，NP完全问题的搜索空间随问题规模呈指数增长，在多项式时间内无法找到最优解。因此许多人尝试采用近似方法，如分批加权排课模型和基于二分图模型的排课系统等求解该问题，然而这些方法在课表信息量比较大的情况下，效率不高。

问题假设

为方便建立面试学生与老师之间合理的分配模型，在实际建模过程中对如下问题进行假设。

1）为保证面试的相对公平性，Y2约束为模型必须满足的约束条件
2）每位老师面试的学生数量尽量均衡也视为硬约束条件，即必须满足的约束，其优先级别略低于Y2
3）Y4：被任意两位导师面试的两个学生集合中出现学生的人数最少，由于保证了老师面试学生数量的均衡性，因此可以将Y4约束优先级别设为最低

由以上假设可得出约束条件优先级别的先后顺序为$\{Y2,Y1,Y3,Y4\}$，在模型优化考虑时，按此优先级别对模型进行优化。

问题建模

定义两个论域学生$S=\{S_1\ S_2\ S_3\cdots S_N\}$，面试老师论域$T=\{T_1\ T_2\ T_3\cdots T_M\}$，模型最终求解是建立f：$N\to M$映射关系

问题的一个可行解就是一张填满的满足一定约束的二维表Table[N^*,M]，N^*表示每个“面试组”成员老师的个数。

面试问题所要优化的对象可以看作一个五元组：

INTERVIEWOBJ＝（α，β，γ，η，ξ）

其中：α：每位老师面试的学生数量均衡

β：面试不同考生的“面试组”成员不能完全相同

γ：两个考生的“面试组”中有两位老师相同数量

η：两个考生的“面试组”中有三位老师相同数量

ξ：任意两位老师的两个学生集合中出现相同学生数超过θ的数量

注：θ表示不能容忍的老师面试结合中相同数量学生的最小值

以上统计是针对一个可行解。

将此问题可以抽象转换成 TTP 求解问题

其定义含义与 TTP 稍有不同

可行解表示：一个二维关系Table[N^*，M]，如果满足约束集合$\{Y2,Y1,Y3,Y4\}$

$$\begin{pmatrix} t_{11} & t_{12}\cdots & t_{1N} \\ t_{21} & t_{22}\cdots & t_{2N} \\ t_{31} & t_{32}\cdots & t_{3N} \\ t_{41} & t_{42}\cdots t_{4N} & \end{pmatrix}$$

其中$t_{ij}\in T=\{T_1,T_2,\cdots,T_M\}$，

R_i表示学生 i 所有面试老师的集合

$R_i \cap R_j = \varnothing$ 或者 $\left|R_i \cap R_j\right| \le 2$，$i = 1,2,\cdots N$，$j = 1,2,\cdots N$

$t_{ik} \ne t_{jk}, k = 1,2,\cdots,N; i,j = 1,2,3,4$ 表示任意一个学生的“面试组”不会出现相同成员。

$\min\left|Interviews_i \cap Interview_j\right|$ $Interviews_i$ 表示老师 i 面试学生的集合，任意两个老师面试学生集合的交集应该最小。

模型求解

遗传算法是一种可以用复杂系统优化计算的鲁棒搜索算法。实践证明，遗传算法对于优化组合中 NP 完全问题非常有效。已有人尝试利用遗传算法求解课表问题（**TTP**），本论文将面试老师分配模型抽象成二维 TTP 问题，因此可以采用遗传算法解决此问题

基本遗传算法可以定义为一个 8 元组：

SGA=$(C,E,P_0,M,\Phi,\Gamma,\Psi,T)$

式中：C————个体编码方式

E————个体适应度评价函数

P_0————初始群体

M————群体大小

Φ————选择算子

Γ————交叉算子

Ψ————变异算子

T————遗传终止条件，一般为迭代次数

与遗传算法性能相关的参数还有两个：

P_c————交叉概率，一般取 0.4~0.9

P_m————变异概率，一般为 0.0001~0.1

基因编码：

编码方式不仅决定个体的染色体排列形式，而且决定个体从搜索空间的基因型变换到解空间表现型时的解码方法。编码方法对交叉算子，变异算子等遗传算子的运算方法及程序实现的复杂度也有影响。

论文采用的是二维矩阵编码方式，可方便利用二维数组保存编码信息，编码解码直观，程序实现复杂度低，进行交叉变异时方便进行约束条件检测并且可以直接使用矩阵变换优化结果。

初始群体:

群体 M 的大小表示群体所含个体的数量，当 M 取值较小时，可以提高遗传速度，但降低了群体的多样性，易引起算法早熟，而当 M 取值较大时，又会降

低算法的运行效率。一般建议 M 取值范围为 20~100。这里取 M=30。

产生初始种群 P_0 时，进行约束检查，除去不满足 Y2 的初始个体，从而使生成的初始种群满足约束条件 Y2。

个人适应度评价函数：

$$F=\sum_{i=1}^{5} W_i \mathrm{InterviewOBJ}[i]$$

W_i 表示第 i 个约束条件相应的选择，权值的选择影响适应度函数，进而会影响最后最优解

1) 约束条件 Y1 对应的权值，

对于一个个体，统计每个老师面试学生的个数

$\bar{X}_i$ 表示第 i 个体所有老师面试学生个数的平均值,利用其方差反映均衡性

X_i 表示老师 T_i 面试学生的个数

$\mathrm{InterviewOBJ}[1]=1/\sum_{i=1}^{m}(X_i-\bar{X})^2$，假定 $W_1=500$；

2）约束条件 Y2，可以在生成种群中，交叉算子，选择算子过程中考虑其约束，将不满足约束条件 2 的个体，从生成种群中剔除，同时在种群在遗传过程中保持这一约束特性，从而减小了遗传搜索空间，提高了搜索效率和精度。

3) 约束条件 Y3，InterviewOBJ[2], InterviewOBJ[3]值为在个体中出现有两个面试老师相同的情况数的倒数，并假定 $W_2=300$， $W_3=20$，由于在实际情况中出现有两个老师相同的情况概率比出现三个老师相同的高，因此设定较高的权值。

4）对于约束条件 Y4，InterviewOBJ[4]值为任意两位老师面试的两个学生集合中出现相同学生数的人数尽量的少

交叉算子：

在群体中随机选择两个染色体作为双亲，将它们中的单元随机交换一部分，也可以利用部分匹配交叉法（Par tally Matched Crossover, PMX）进行交叉运算。

$$\begin{pmatrix} 280 & 418 & 281 & 416 \\ 416 & 0 & 416 & 0 \\ 283 & 513 & 417 & 527 \end{pmatrix} \qquad \begin{pmatrix} 293 & 285 & 284 & 286 \\ 286 & 289 & 280 & 289 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 280 & 418 & 281 & 416 \\ 286 & 289 & 416 & 0 \\ 0 & 0 & 417 & 527 \end{pmatrix} \qquad \begin{pmatrix} 293 & 285 & 284 & 286 \\ 416 & 0 & 280 & 289 \\ 283 & 513 & 0 & 0 \end{pmatrix}$$

选择算子：

遗传算法使用选择算子来对群体中的个体进行优胜劣汰操作，选择算子操作的主要目的是为了避免基因损失，提高全局收敛性和计算效率。

在种群中选择个体进入下一代种群，首先采用的策略是将适应度数值最高的个体，即保留最优基因个体遗传到下一代种群中，其余 M－1 个个体采用轮盘赌选择算法如：

$$\frac{F[i]}{\sum_{i=1}^{M} F[i]}$$

具有较大适应度函数值的个体较容易被选中，同时也保证了一定的随机性，使下一代种群在保证优良基因的同时，也兼顾基因的多样性。

变异算子：

变异算子是产生新个体的辅助方法，它决定了遗传算法的局部搜索能力。变异算子有助于避免遗传算法过早收敛。根据论文提到的编码方式，采用一种日变异的变异方法。

变异算子实现如下（对种群内每一个体）

第一步：产生 0～1 之间随机数 num1,判断是否需要变异。如果 num1< P_m，执行第二步，否则，保留该个体。

第二步：产生 1～δ 之间随即整数 num2 和 num3,交换 X 坐标分别为 num2 和 num3,交换相同的 Y 坐标。

同时交换完之后要检测是否违反约束 Y2

文理老师分科后情况

对于将文科老师和理科老师分半，可将已有模型修改

定义三个论域：

学生论域 $S=\{S_1,S_2,\cdots S_N\}$，

文科老师论域 $T_{art}=\{T_1,T_2,\cdots T_{m'}\}$，

理科老师论域 $T_{science}=\{T_{m'+1},T_{m'+2},\cdots T_M\}$

$Interviewee_j=\{T_0 \quad T_0' \quad T_1 \quad T_1'\}$ $T_0,T_0'\in T_{art}$， $T_1,T_1'\in T_{science}$， $Interviewee_j$ 表示面试

学生 j 的老师的集合

这样可以在遗传算法编码的过程中可以考虑建立三维编码模型，同时在交叉，选择，变异算子中要考虑 $T_{art}, T_{science}$ 之间不发生冲突。

模型不足和改进方案

论文基本建立了处理学生面试问题的模型,并在一定假设前提下求得了一个相对最优解模型存在的主要不足是优化目标不够精确，遗传搜索中适应度函数选择不够精确,这可能会影响搜索效率和准确性。

详细分配方案及编程代码见附录 A,B

问题 3

新的假设条件下解决问题 1

解决思路与问题 1 相同，先将原问题转化为一个等价命题，然后两步解决等价问题.

对偶命题 2：设 G 是一个 m 阶无向完全图 Km（m 为偶数），其中 $S_1=\{v_1,v_2,...,v_{\frac{m}{2}}\},S_2=\{v_{m/2+1},...,v_m\}$，为 G 的 m 个顶点的集合，n 为由 S_1，S_2 各两个顶点组成的无重复边的 K_4 子图的数目，则当 n= N 时，求能够满足条件的最小的 m 值。

为了解决的这个问题，也是可以通过下面两个步骤：

1.确定 m 的一个较优的下界值。

由于对偶命题 1，对偶命题 2 描述的是两个相象的问题，只是对偶命题 2 中对于 K_4 的约束更加严格，m 的上界区间比命题 1 小，所以命题 1 的两个引理同样可以使用.

引理 1：设 G 是一个 m 阶无向完全图 Km（m 为偶数），其中 $S_1=\{v_1,v_2,...,v_{\frac{m}{2}}\},S_2=\{v_{m/2+1},...,v_m\}$，为 G 的 m 个顶点的集合，n 为由 S_1，S_2 各两个顶点组成的无重复边的 K_4 子图的数目，则必有：

$$m\geq\frac{1+\sqrt{1+48n}}{2}$$

引理 2：

对于一个 m 阶完全图 G，n 表示 G 中无重复边的 K4 的个数，则有：

（1）m mod 3 = 0 时，$m\geq\frac{3+\sqrt{9+48n}}{2}$

（2）m mod 3 = 2 时，$m\geq\frac{2+\sqrt{4+48n}}{2}$

（3）m mod 3 = 1 时，$m\geq\frac{1+\sqrt{1+48n}}{2}$

引理 3：设 G 是一个 m 阶无向完全图 Km（m 为偶数），其中 $S_1=\{v_1,v_2,...,v_{\frac{m}{2}}\},S_2=\{v_{m/2+1},...,v_m\}$，为 G 的 m 个顶点的集合，n 为由 S_1，S_2 各

两个顶点组成的无重复边的 K_4 子图的数目，t 为 S_1 中顶点组成的边在所有的无重复边的 K_4 中的利用率, 则必有：

$$t \le \frac{2(m-1)}{3(m-2)}$$

证明： S_1 的顶点数为 $\frac{m}{2}$，则能够组成的边数为 $C_{m/2}^2$

进一步 $$6t\,C_{m/2}^2 \le C_m^2$$

所以 $$t \le \frac{C_m^2}{6\,C_{m/2}^2}$$

即 $$t \le \frac{2(m-1)}{3(m-2)}$$

引理 4：设 G 是一个 m 阶无向完全图 Km（m 为偶数），其中 $S_1 = \{v_1, v_2, ..., v_{\frac{m}{2}}\}, S_2 = \{v_{m/2+1}, ..., v_m\}$，为 G 的 m 个顶点的集合，n 为由 S_1，S_2 各两个顶点组成的无重复边的 K_4 子图的数目，r 为 G 中除去无重复边的 K_4 子图后边的剩余数目，则必有：

$$r \ge \frac{m^2}{3} - \frac{m}{3}$$

证明：设 t 为 S_1 中顶点组成的边在所有的无重复边的 K_4 中的利用率，由引理 3 可以知 ：

$$t \le \frac{2(m-1)}{3(m-2)}$$

除去无重复边的 K_4 子图后边的剩余由两部分组成，一是 S_1，S_2 内部的点组成的边，共 $2(1-t)\,C_{m/2}^2$；二是 S_1，S_2 之间的点组成的边，共有 $(\frac{m}{2})^2 - 4t\,C_{m/2}^2$，

则 r 可以表示为:

$$r = 2(1-t)\,C_{m/2}^2 + (\frac{m}{2})^2 - 4t\,C_{m/2}^2$$

即有：$r \ge \frac{m^2}{3} - \frac{m}{3}$

由上很容易得到下面的引理

引理 5：设 G 是一个 m 阶无向完全图 Km（m 为偶数），其中

$S_1=\{v_1,v_2,...,v_{\frac{m}{2}}\},S_2=\{v_{m/2+1},...,v_m\}$，为 G 的 m 个顶点的集合，n 为由 S_1， S_2 各两个顶点组成的无重复边的 K_4 子图的数目，则必有：

$$m\geq\frac{1+\sqrt{1+36n}}{2}$$

2 利用 1 中计算的下界，向上搜索，找到最理想的 M 值

同问题 1 的过程,用 0,1 数字矩阵来表示 G, Km.

除了上面那些相同的部分,还需要一个描述老师状态的矩阵 S,仅仅查询. $S=[\begin{matrix}1 & 2 & 3 & 4 & .... & m-1 & m\\ 1 & 1 & 1 & 1 & .... & 0 & 0\end{matrix}]$

利用上面的结论，在给定 n 的条件下，可以获取对应 m 的一个下界，从此时的 m 开始，不妨记为 M = m。

与问题 1 K_4 搜索算法相比,差别在于检查一个四点组合时,首先检查其在 S 矩阵中的状态是否满足 1,1,0,0 的组合,如果满足,再继续检查其他的条件（同问题 1);如果不满足,则选择下一组 A[i][s]，A[i][t];如果所有的组合均不成立,则将该考察点变为 0,继续选择下一考察点,直到将 A 变成零矩阵.除此,与问题 1 处理过程完全一致.

问题 4

1. 新分配方案提出的背景。

正如我们所看到的，在高校自主招生的具体实施过程中，如何保证考生与面试老师之间分配的均匀性以及面试工作的公平性和信度成为社会公众和学术界关注的焦点问题。学生面试作为一项主观性很强的评价工作，考生与面试老师之间分配的均匀性是实现面试工作的公平性和信度的基础和保证，面试的公平性又是考生与面试老师之间分配均匀性的具体体现，两者是相辅相成的统一体，任何一方都不能偏废，否则就会出现过度追求分配的均匀性反而无法保证面试效果实际公平等极端结果的出现。

鉴于面试老师的专业背景和研究领域有很大不同、老师的评分习惯、提问方式、提问内容乃至个人喜好等也会有较大差异，以及考生的学科背景、临场发挥、乃至个人性格因素等的差异性都会对评价结果产生不同程度的影响。例如，面试中经常存在“晕轮效应”或“触角效应” 的影响。面试老师很可能把自己对考生的判断建立在考生的第一印象、高中所在学校等因素的基础上，而相对忽视了对考生学习和实践能力等方面的考察。再如，由于面试为每位学生提供的面试时间和空间都非常有限，很多考生无法将真实的自己充分展示给面试老师，而具有表演天分、性格外向的同学则很容易赢得面试老师的好评。再例如，在现实面试过程中，面试老师中往往存在一位或几位德高望众的老师，他或她虽然并不直接干涉其他面试老师的评分，但其在面试中的态度和倾向性会对其他面试老师具有相当大的导向作用，结果造成一个人的意见代表了其他面试老师的意见。种种情形在此不一而足。

2 新分配方案的内容和公平性测评方法。

为了保证面试的均匀性、公平性和信度，特别是鉴于当前学科的交叉性日益加强，学生的综合素质越来越成为人才评价和社会需求的重要因素，我们提出以下方案：

某高校在今年的自主招生中，经过初试合格进入面试的考生有 N 人，拟聘请老师 M 人。每 4 位老师同时面试 3 个学生，每位学生随机参加 3 场面试，要求如下：

（1）每位老师面试的同学数量应尽量均衡；

（2）同一考生在 3 轮面试中的老师尽量完全不同；

（3）同一考生在 3 轮面试中出现两位或三位队友相同的情形尽量的少；

（4）一轮面试中三名面试学生的文、理科背景不完全相同；

（5）在一个面试过程中要求三位考生对同一问题首先分别提出自己的观点，然后以辩论等形式论证自己的观点，最后归纳自己和队友的观点，也可提出相关解决措施或构想。

针对上述分配方案的公平性测评方法。

在有众多主观评定因素的面试情境中，面试老师之间的差异性是产生评价结果误差，影响面试公平性和信度的重要原因之一。我们可以采用以下方法对面试公平性进行测评:

（1） 运用肯达尔一致性系数（W 系数）进行定量测评。我们可以根据每位面试老师给每位考生的面试分数的排列顺序计算肯达尔一致性系数。若该系数的值如果超过 0.8，则可以认为各位面试老师的评价结果具有显著的一致性，则面试的公平性和信度就高。不过此种方法仅仅根据等级排序的数据，很可能会造成数据的缺失。

（2） 运用概化系数进行定量测评。我们可以用方查分量来估计概化全域（例如可以增减面试老师人数）下概化系数的变化，来找到有效控制误差，提高面试公平性和信度的最优面试方案。

（3） 根据面试分数的统计结果，分析参加面试考生的分数排序是否与各位面试老师根据总体印象的排序一致。从理论上讲，一个公平性和信度较高的面试，各位面试老师的意见应当大体一致，各位面试老师按主观印象进行的排序结果与按分数的排序结果也应当基本一致。

3 与原方案相比，该分配方案有以下优点：

（1）老师的工作量和老师人数并没有增加。此方案中一轮面试的 4 位老师同时为 3 位学生打分，每位学生参加 3 次面试，这样并没有增加面试老师的人数和老师的工作量。

（2）学生面试的公平性和信度大大增强。每个学生参加三轮面试，共有 12 位面试老师为同一个学生打分。这样评分结果受单一老师的专业背景、个人喜好

和评分习惯等主观因素的影响较少。

（3）此种方案为较为全面地评价学生的综合素质和学习潜能提供了一种思路。该方案要求三个学生对同一问题进行阐发。在此过程中，学生提出问题、分析问题、解决问题以及归纳观点的能力、沟通能力等都得到了较为充分和全面的展示。

（4）此种方案更顺应当前学科交叉的大背景。该方案要求一轮面试中三名面试学生的文、理科背景不完全相同，这样对同一个问题大家的出发点、思路就不尽相同，在相互阐发的过程中就可以互相借鉴、为我所用，而不局限于自己的学科领域内，故步自封。正如我们所看到的，当前学科之间的差异性越来越小，学科间相互交叉、融合已经成为大的发展趋势。人文社会科学的研究和自然科学的探讨和应用越来越紧密地联系起来，正如这道数学建模题目所反映地一样，高考自主招生这个社会问题与数学方法和计算机工具完美地结合在一起。

（5）此种方案更符合大学生培养的客观实际。当前大学生的教育和培养远不是传授书本知识那样简单，更重要的是培养学生的学习研究能力和团队合作意识。而该方案恰恰反映了学生对问题进行整理归纳形成个人观点的能力以及与人沟通和合作的能力。

4 该分配方案的不足和仍需改进的地方。

（1）学生的分组较为繁难，计算量大。该方案中每 4 位老师要同时面试 3 位学生，每位学生随机参加 3 场面试，这样于与原方案相比，学生的分组问题考虑的因素大大增加，计算的难度加大。

（2）该方案的相对公平性。尽管这种新方案的提出在一定程度上使面试的公平性和信度得到增强，然而无论如何，面试毕竟是一个主观性很强的评价工作，老师的专业背景和研究领域、提问内容、提问方式、评分习惯和个人喜好以及学生的学科背景、临场发挥、乃至个人性格因素等都会对评分结果产生或大或小的影响。因此，在该方案的基础上，仍需探索综合考量各种影响因素，并对其分别予以量化的途径。例如，可以根据不同影响因素对面试公平性的影响度大小，对各种因素赋予不同的权重，然后对评分结果进行加权平均，来得到更为公平的评价结果。

参考文献:

[1] 陶世群 蒲保兴 基于遗传算法的多级目标非平衡指派问题求解 系统工程理论与实践 2004.8
[2] 郑月锋 黄德才 刘瑞阳 遗传算法在求解时间表问题中的应用 浙江工业大学学报 2006.4
[3] 业宁 梁作朋 董逸生 一种遗传算法的 TTP 问题求解方法 东南大学学报 2003.1 33 卷第 1 期
[4] 张威 陈莘萌 用遗传算法解决条件约束问题的一种新思路及其在 TTP 问题中的应用 计算机工程与应用 2002.21