基于图论的高校排课系统实现*

张　健
(重庆师范大学 数学与计算机科学学院，重庆 400047)

摘　要：在参照多种排课算法后，对排课资源进行合理抽象并建模，利用图论染色和最优匹配原理来简化算法，并结合数据库技术给出排课问题的数据存储模式，从而提出一套完整而实用的高校排课系统的可行实施方案。

关键词：二分图匹配；图论染色；排课算法
中图分类号：TP392；O157.6　　**文献标识码**：A　　**文章编号**：1672-6693(2005)01-0035-04

Implementation of University Course Timetabling on Graph Theory

ZHANG Jian
(College of Mathematics and Computer Science, Chonqqing Normal University, Chongqing 400047, China)

Abstract: After consulted with manifold algorithm of arranging course, abstracted logically and molded to resource of arranging course, in addition to utilizing coloring and optimization matching of digraph in graph theory to predigest algorithm and combining date-base technology to present data storage mode of arranging course, this paper provides a full and feasible timetabling solution of arranging course in universities.

Key words: digraph matching; graph coloring; timetabling arithmetic

近 20 年来，已经有许多学者针对不同应用环境的排课问题给出不同的解决方案。通过图论的方式来研究排课问题是一个比较典型的方法，但多数情况下对排课的数学抽象都过于简单[1,2]。

另外，近段时间也出现过利用人工智能如遗传算法[3]、专家系统[4]等方法对排课问题在理论上作近似优化讨论，但其实验数据较少，其搜索算法或推理过程对于这样已被证明为 NP 难的问题，无法从数学角度给出这些方案是否为有效性的证明。特别来说，有些算法，如遗传算法，还可能逐渐偏离全局最优解和不稳定性[5]。

总之，许多方案对课表的优化角度不尽相同，有的甚至没有优化标准，开发软件时可操作性困难。事实上排课是多约束条件下的多目标优化问题，另外计算机代数系统对图论模型给出了一些计算[6]，可以为图论模型的有效性做出一定的回答。故本文从图论角度对排课问题重新抽象并结合数据库技术给出问题的数据存储模式，最后提出了一种切实可行的排课算法，从而给出一套完整而实用的高校排课系统的可行方案。

1　排课问题

在排课问题中，其主要任务是将具有多种属性的各种资源，如教师、教室、班级、学生、课程、时间等，以一个周期的方式进行合理地匹配，使其不发生冲突，其数学模型如下。

假设排课资源的每个实例都是一个多属性的元组，其元组集用大写字母记为：TE(教师)、CL(班级)、ST(学生)、RO(教室)、TI(时间)、CO(课程)、PL(计划或称为排课目标)，相应的实例用小写字母表达。设 $x_n = (te_{n1} \cdots te_{ni}, cl_{n1} \cdots cl_{nj}, st_{n1} \cdots st_{nk}, ro_{n1} \cdots ro_{nr}, ti_{n1} \cdots ti_{nr}, zc_n, co_n)$ 是对计划为 pl_n 的一个解，其中 zc_n 是 x_n 中占用周次。这里说排课资源进行了合理分配有以下两层含义。

其一是课程安排至少要满足以下硬性约束，即是本次目标规划一个可行解。

* 收稿日期：2004-10-08
作者简介：张健(1972-)，男，成都人，讲师，主要研究方向为凸分析与多目标规划及最优算法。

(1) $zc_n \wedge zc_m > 0$ 且 $te_{ni} = te_{mi}$, $\Rightarrow ti_{ni} <> ti_{mj}$,即:同一教师必须在不同时间;

(2) $zc_n \wedge zc_m > 0$ 且 $cl_{ni} = cl_{mi}$, $ti_{ni} <> ti_{mj}$,即:同一班级必须在不同时间;

(3) $zc_n \wedge zc_m > 0$ 且 $st_{ni} = st_{mi}$, $ti_{ni} <> ti_{mj}$,即:同一学生必须在不同时间;

(4) $zc_n \wedge zc_m > 0$ 且如果 $st_{ni} \in cl_{mj}$,则 $ti_{ni} <> ti_{mj}$,即:个人课表服从集体课表;

(5) $zc_n \wedge zc_m > 0$ 且 $ti_{ni} = ti_{mj}$ 且 $ro_{ni} = ro_{mj}$ 且 ro_{ni}.容纳类型 =“单计划”$\Rightarrow n = m$,

即:同一时间同一教室只能排一个计划;

(6) ro_{ni}.容量 $\geqslant p(x_n)$,其中函数 $p(x_n)$ 表示计划为 n 的上课人数;

(7) ro_{ni}.类型 $= t(x_n)$ 或者 ro_{ni}.兼容类型 $= t(x_n)$,其中函数 $t(x_n)$ 表示计划为 n 所需要的教室类型;

(8) $\sum ti_{ni}$.课时 $> x_n$.课时,即指定的教学时间能够完成该课程的教学任务。

其二是课程安排还应有优化目标,让课表具有向最优策略滑动的趋势。

(1) $\mathrm{DE} = \mathrm{Max} \sum (d(ti_{ni}) - d(ti_{nj}))^2$,其中函数 $d(ti)$ 返回 ti 所在星期数,DE 代表同门课程的分布程度,即同一门课程尽量不在同一天上课;

(2) $\Delta_k = \mathrm{Min} \sum (k_n(st_i) - a(st_i))^2$,其中 $k_n(st_i)$ 返回 st_i 星期 n 的课时总数,$a(st_i)$ 返回一星期为周期,每天的平均课时数。即课程安排尽量均匀;

(3) $\Delta_t = \mathrm{Max} \sum ti_{ni}$.优先级 $* p(x_n)$,即让最广泛的同学享受到最好的听课时间。

(4) $\Delta_d = \mathrm{Min} \sum dxy(st_i, ti_n, ti_{n+1}) + \sum dxy(te_i, ti_n, ti_{n+1})$,其中 dxy 代表学生或者教师在时间由 ti_n 变动 ti_{n+1} 教室变化距离,即课表安排针对人来说,相对集中,避免上下课教学楼通道过分拥挤,并且学生或教师能够及时赶到下堂课程的教学地点。

(5) $\Delta_r = \mathrm{Min} \sum (ro_{ni}$.容量 $- p(x_n))$。即教室利用率评价标准,避免超大教室安排小班级的教学。

(6) 其它优化指标。如教师预定时间的满足度,最佳上课时间利用度等。

2 相关的图论原理

引理2.1 求解图G的色数是一个NP问题[7]。

目前已知最好求图色数算法(对图的结点而言)具有指数的最坏情形时间负复杂性,即使求图近似的色数的近似值的问题也是困难的。已经证明,假设存在具有多项式最坏情形时间复杂性的算法,图的色数可以达到2倍近似值,那么也存在具有多项式最坏情形时间复杂度的求图色数的算法[7]。近年也出现利用遗传算法来对图着色近似求解[8]。

算法2.2 (Welch Powell 近似着色法[9])设 $G = <V, E>$, $V = \{v_1, v_2, \cdots v_n\}$, $E = \{e_1, e_2, \cdots e_m\}$,记 $N(v_i)$ 表示与 v_1 相邻的结点集,$\lambda(v_i) = c$ 表示 v_i 所着颜色为 c,则算法如下:

(1) 将图 G 中的结点按照度数递减排序,并令 $\lambda(v_i) = 0$(对所有的 v_i)

(2) $i := 1$

(3) $c := 1$

(4) 对任意 $v \in N(v_i)$, $\lambda(v) \neq c$,且 $\lambda(v_i) = 0$,则 $\lambda(v_i) = c$,并转入到(6)步

(5) $c := c + 1$

(6) 若 $i < n$,则 $i := i + 1$,并转入到(3)步。

定理2.3 设图 G 结点的最大度数为 Δ,那么由算法2.2产生一个 $\Delta + 1$ 的着色,因此有 $x(G) \leqslant \Delta + 1$[2]。

定义2.4 (最优匹配)设图 G 是无环图,M 是边集 $E(G)$ 的子集,如果 M 中任意两条边在图中均不相邻,称 M 是图的一个匹配。对赋权完全二分图 $G = (K_{n,m}, w)$ 求解一个总权数最大的匹配,称为最优匹配。

Kuhn(1955)和 Munkres(1957)提出对 $(K_{n,n}, w)$ 中最优的有效算法[2],此处不再赘述。

3 排课问题的图论模型

由本文第1部分对排课问题的讨论,可以将排课问题看成两个子问题,即:(1)时间的指定,(2)教室的指定。

3.1 时间的指定

通过分析排课可行解的限制条件(1)~(4)知道,造成开课计划冲突的因素,主要是两个计划中具有共同的学生或者教师或者班级,所以将学生、教师和班级看成一个因素,定义成参与者。另外,将开课计划看成图 G 中点集 V,并将具有共同参与者的两个开课计划 v_i 和 v_j 连接成边,则对开课计划指定时间就相当于为每个点指定颜色,所以该问题即为图论着色问题。在着色过程中并尽量满足优化条件中的(1)~(3)。

实际上由引理2.1知这是一个难以解决的问题。考虑到课表将来调课的灵活性和可扩充性,该

问题不应该过分追求图的色数。因此,在工程上往往采用 Welch Powell 近似着色法,并限制了每种颜色的点集数,防止某种颜色的点过多,导致无法安排教室。事实上 Welch Powell 着色法同手工排课操作类似,即先安排合班数多的计划,再安排一般要求的计划。当然,该算法也仅仅类似专家系统,是一个启发式算法,但算法比遗传算法和专家系统简单且具有稳定性和可靠性,工程上容易实现。

另外,由排课问题形成的图各个点的度数都不是太大,根据 3 年 6 学期对我校排课计划数据的统计得出最大度数(即课程冲突量)不超过 100,平均度数为 12。由定理 2.3,排课时间分配的可行解几乎是存在的,否则人工辅助进行干预。

一般地,对多于 2 000 个计划的高校排课采用 Welch Powell 近似着色法时,最小颜色数同颜色的点集最饱满,后续具有相同颜色的点数迅速减少。所以在为计划指定时间时,按照时间的重要性从最小颜色号依次指定时间,并要求满足可行解约束条件(8)。

3.2　教室的安排

当为开课计划指定了时间后,剩下的工作就是指定教室。设 V_1 具有相同的时间计划点集,并用点集 V_2 表示所有的教室,再通过排课可行解约束条件(5)~(7)将凡是 V_1 中点可以安排到 V_2 中的点连接成边,最后通过最优条件(5)为所有的边赋权,则形成一个赋权二分图。这样,教室的安排就可以考虑成图论的最优匹配问题,然后用 Kuhn-Munkres 有效算法进行求解最优值。

4　排课系统方案设计

根据以上分析,本文从三个方面给出一个易于实施的排课系统设计方案。

4.1　代码设计

首先,由于排课中的数据具有相关性,建议每个实体的关键字信息代码都附加一个检验码,防止数据的录入错误。其次,由 3.1 中的分析,最好将教师、学生、班级进行统一编码。例如将教师和班级都统一编成 n 位码,具体含义由读者定义,而学生编码 = 班级码 + 三位序号码,则为以后检测时间冲突带来方便。最后,关于占用周次的编码,建议用 4 个字节 32 个位表达,来代表一学期中 32 周哪些周次将被占用。

4.2　数据库设计

排课系统的数据库主要有教师、学生、班级、教室、时间、课程等固定信息表。另外,其它信息表及属性如下。

教学计划:计划编号、课程编号、占用周次、周学时、周课次、等级、教室类型、教室容纳类型、预定时间、预定教室。

开课计划:开课编号、计划编号、学时、时间编号、教室编号、冲突量。

选课信息表:参与人编号、计划编号。

时间冲突表:开课编号 1、开课编号 2(即用于存储图的边)。

时间教室表:时间编号、教室编号、开课编号 1、…、开课编号 32。

4.3　过程设计

4.3.1 数据输入过程设计　排课系统的数据输入主要就是教学计划的输入,在输入教学计划时可以同时输入该计划的参与人,将这些信息分别存储到教学计划表和选课信息表。当然,该操作可以在网络上同时进行,从而做成一个选课模块。

4.3.2 排课数据预处理　首先,排课数据的预处理就是将每个教学计划拆分成适合时间片的若干个实际的开课计划。例如,假设给出的每个排课时间片段可以安排 2 个学时或者 3 个学时,则每周行课 5 学时的计划,将拆分成 2 学时和 3 学时的形成两个正式开课计划,形成着色图的点。

其次,将对开课计划进行两两检查,凡是有时间冲突的开课计划则记录到时间冲突表中(建立着色图的边)。假设班级编码为 n 位,开课计划中凡是具有相同前 n 编码的参与者,则表示冲突。将各开课计划的冲突量填写到开课计划表中。此操作可以用极少数的 SQL-查询语言完成。

时间表和教室表作笛卡儿积,写入时间教室表,置所有开课编号为零。

4.3.3 排课算法　排课可以分为自动排课和人工辅助排课。此处给出的算法为自动排课算法。

(1)对凡是具有预定时间的开课计划进行扫描,为这些开课计划指定预定时间,并检测时间是否有冲突。即检查相同时间的开课计划中在冲突表中是否存在边,如果存在则取消一个开课计划的预定时间。检查冲突后在按照最优原则指定教室或指定预定教室;

(2)将时间号按照优先级排序,记为 $t(1)$,$t(2)$,…$t(n)$;

(3) $c := 1$;

(4)在开课计划表中按照条件"开课计划.课时 $<= t(c)$.课时 and ($t(c)$.时间编号 $=0$ or $t(c)$.时间编号 $=0$)"选出开课计划表,并按照时间编号降序、等级升序、冲突量降序进行排列,记为 $k(1)$,$k(2)$,…$k(m_c)$;

(5) $i := 1$;

(6) $k(i)$.时间编号 $=0$,对任意小于 i 且有 $k(j)$.时间编号 $=t(c)$ 的 j,有 $k(j)$ 与 $k(i)$ 不构成边,则 $k(i)$.时间编号 $=t(c)$;

(7)若 $k(i)$ 着色成功,则进行教室最优匹配。若匹配不成功,则 $k(i)$.时间编号 $=0$;

(8) $i := i+1$,如果 $i > m_c$ 转入(8),否则转向(5);

(9) $c := c+1$,如果 $c <= n$ 转入(3),否则停止。

一般地,在做人工辅助排课或调课时,可采用算法中的(5)和(6)为操作员提供决策信息。

5 结束语

按照上述思想,作者设计与开发了高等院校排课系统,并经重庆师范大学2001年与2002年的排课数据测试成功,于2003年重庆师范大学正式投入使用,经3学期运行实践来看,从资源利用率、学生上下课转移量、排课速度等方面都取得了满意的排课效果。同时排课系统的解决可为各高校以课程排课为中心的选课系统、成绩管理、教学检查等教务管理信息化建设铺平道路。另外,本文从多角度为排课问题给出数学抽象,希望能够为数学爱好者提供进一步研究的素材。

参考文献:

[1] WERRA D. Restricted Coloring Models for Timetabling Discrete Mathematics [J]. 1997,1656-1666(15):161-170.

[2] 殷剑宏,吴开亚.图论及其算法[M].合肥:中国科学技术大学出版社,2004.152.

[3] 唐勇,唐雪飞,王玲.基于遗传算法的排课系统[J].计算机应用,2002(10):93-94.

[4] SOLOTOREVSKY G,GUDES E,MEISELS A. RAPS:a Rule-based Language for Specifying Resource Allocation and Time-tabling Problems Knowledge and Data Engineering[J].1994,6(5):681-697.

[5] 阎平凡,张长水.人工神经网络与模拟进化计算[M].北京:清华大学出版社,2000.386.

[6] 李世奇.基于Maple的图的构建和计算[J].重庆师范学院学报(自然科学版),2002(4):35-39.

[7] KENNETH H ROSEN.离散数学及其应用(第4版)[M].北京:机械工业出版社,2002.495.

[8] 刘西奎,李艳,许进.基于遗传算法的图关联着色算法[J].工程数学学报,2004(1):41-47.

[9] 左孝凌,李为鑑,刘永才.离散数学[M].上海:上海科学技术文献出版社,1982.

(责任编辑 李若溪)