An approach for the Class/Teacher Timetabling Problem using Graph Coloring

G. S. Bello · **M. C. Rangel** · **M. C. S. Boeres**

Received: date / Accepted: date

Keywords Timetabling · Graph Coloring · Metaheuristics · Tabu Search

1 Introduction

This work addresses both the Class/Teacher Timetabling Problem (CTTP) and the Graph Coloring Problem (GCP). In the case of the Class/Teacher Timetabling Problem the specification with both hard (which must be satisfied) and soft constraints (which should be satisfied) was chosen. This formulation, proposed in (Santos et al 2007; Souza et al 2003; Souza 2000), is adapted to instances representative of Brazilian schools. In the next section, the CTTP description is presented. Section 3 shows the approach proposed in this work and Section 4 is dedicated to computacional results and conclusions remarks.

2 The Class/Teacher Timetabling Problem (CTTP) Description

The CTTP builds the weekly scheduling of teachers and classes. It is described below, adapted to the format and nomenclature defined by Neufeld and Tartar in (Neufeld and Tartar 1974). Given

- A set of teachers $T = \{t_i\}$, $i = 1, \dots, \alpha$;
- A set of classes $C = \{c_j\}$, $j = 1, \dots, \beta$;
- A set of weekly hours $H = \{h_k\}$, $k = 1, \dots, \sigma$, in which λ week days with μ daily periods, define $\sigma = \lambda \cdot \mu$ different hours;
- A $\alpha \times \beta$ requirements matrix $R = [r_{ij}]$, where $r_{ij} \geq 0$ and r_{ij} is equal to the number of weekly lesson hours of teacher t_i for class c_j;
- A $\alpha \times \sigma$ teachers unavailability matrix $D = [d_{ik}]$, with $d_{ik} = 1$, if teacher t_i is unavailable at hour h_k and $d_{ik} = 0$, otherwise.

Universidade Federal do Espírito Santo
Av. Fernando Ferrari, 514, Campus de Goiabeiras, Vitória, ES, Brazil, 29075-910
Tel.: +155-27-33352127
Fax: +155-27-33352850
E-mail: geraldo.bello@terra.com.br, {crangel,boeres}@inf.ufes.br

- A $\alpha \times \beta$ daily lessons hours limits matrix $U = [u_{ij}]$, where $0 \le u_{ij} \le 2$ and u_{ij} is equal to the maximum number of daily lesson hours of teacher t_i for class c_j;
- A $\alpha \times \beta$ double lessons matrix $S = [s_{ij}]$, where $s_{ij} \ge 0$ and s_{ij} is equal to the minimum required number of double lessons (lessons scheduled in two consecutive periods on the same day) required of teacher t_i for class c_j;

The ρ^{th} meeting of teacher t_i and class c_j is denoted by m_{ij}^{ρ}, where $1 \le \rho \le r_{ij}$. The r_{ij} meetings of teacher t_i and class c_j is represented by $M_{ij} = \{m_{ij}^1, \ldots, m_{ij}^{r_{ij}}\}$. The $\sum_{j=1}^{\beta} r_{ij}$ meetings of teacher t_i with the classes $c_j \in C$ is represented by the set $M_i = \bigcup_{j=1}^{\beta} M_{ij}$.

A CTTP solution must satisfy the following hard (1 to 5) and soft (6 to 8) constraints:

1. Every teacher can not be allocated to more than one lesson in the same hour;
2. Every class can not be allocated to more than one lesson in the same hour;
3. Every teacher can not be allocated at hours that they are unavailable;
4. All teachers must fulfill their weekly workload;
5. Every class can not have more than two daily lesson hours with the same teacher;
6. The lessons scheduled for each teacher should be concentrated in the least possible number of days;
7. Double lessons required by teachers should be satisfied whenever possible;
8. In the time schedule of teachers, the occurrence of hours without activity (gaps) between two lesson hours in the same day should be avoided.

A CTTP solution is evaluated by the objective function (Santos et al 2007; Souza et al 2003)

$$\min f(Q) = \omega \cdot f_1(Q) + \delta \cdot f_2(Q) + \pi \cdot f_3(Q) \tag{1}$$

with

$$f_3(Q) = \sum_{i=1}^{\alpha} \theta \cdot g_i + \varphi \cdot v_i + \psi \cdot l_i \tag{2}$$

where (1) is constituted by $f_1(Q)$, $f_2(Q)$, $f_3(Q)$ and their respective weights ω, δ and π. The $\alpha \times \sigma$ matrix $Q = [q_{ik}]$ corresponds to a CTTP solution. The value $q_{ik} \in \{-1, 0, 1, \ldots, \beta\}$ where $q_{ik} = j$, if the teacher t_i teaches the class c_j at hour h_k, $q_{ik} = 0$, if the teacher t_i is available for allocation at hour h_k and $q_{ik} = -1$, if the teacher t_i is unavailable at hour h_k (it corresponds to $d_{ik} = 1$ in the matrix $D = [d_{ik}]$). The value $f_1(Q)$ represents the number of occurrences that a class c_j is allocated to more than one lesson at a hour h_k, $f_2(Q)$ represents the number of allocations that exceed the daily maximum number of lessons u_{ij} of the teacher t_i to the class c_j and $f_3(Q)$ measures the requests from teachers which are not satisfied. Moreover, in (2) the value $f_3(Q)$ is composed by g_i, v_i and l_i and their respective weights θ, φ and ψ. The term g_i represents the number of gaps in the scheduling of lessons of each teacher t_i, v_i is the number of weekdays each teacher t_i needs to come to school for teaching and l_i is the non-negative difference between the minimum number of double lessons, given by $\sum_{j=1}^{\beta} s_{ij}$, required by each teacher t_i and the effective number of double lessons scheduled for teacher t_i.

The $f_1(Q)$ and $f_2(Q)$ components, in the objective function, evaluate the feasibility of the hard constraints, while $f_3(Q)$ evaluates the feasibility of the soft ones. For this reason the weights ω, δ and π are chosen so that $\omega \gg \delta \gg \pi$. Similarly, the weights θ, φ and ψ are chosen to reflect the relative importance of the components g_i, v_i and l_i respectively.

3 The approach proposed

The basic version of CTTP (Neufeld and Tartar 1974) can be transformed into a graph coloring problem, according to the following correspondences:

- Each vertex of the graph represents a lesson;
- An edge joining two vertices indicates that the respective associated lessons can not be scheduled at the same hour;
- Each color represents one hour of the corresponding timetable.

The graph $G = (V,E)$ associated with the CTTP is called **adjunct graph**. The vertex $v \in V$ that corresponds to meeting $m^{\rho}_{ij} \in M_{ij}$ is denoted by v^{ρ}_{ij}. The subset in V that corresponds to M_i is denoted by V_i.

Each unavailability constraint of CTTP (determined by the unavailabilities matrix D) corresponds to a condition that some vertices of the adjunct graph G can not be assigned to a particular color (Neufeld and Tartar 1974).

An adjunct graph G with prevention of color assignment constraints can be converted to an adjunct graph G'' without prevention of color assignment constraints, such as G is σ-colorable if and only if G'' is σ-colorable (Neufeld and Tartar 1975) .

The adjunct graph G'' can be colored using a Tabu Search algorithm for graph coloring called **Tabucol** (Hertz and de Werra 1987). The objective function to be minimized used in this algorithm is $\bar{f}(c) = \sum_{k=1}^{\sigma} |E(C_k)|$, where c is a solution such that $E(C_k)$ is the set of edges with both terminal vertices in the color class C_k, $k = 1,\ldots,\sigma$.

A solution $c = (C_1,\ldots,C_\sigma)$ such that $\bar{f}(c) = 0$ (an σ-coloring of G''), corresponds to the minimization of the $f_1(Q)$ component of the CTTP objective function $f(Q)$.

This work proposes the following adaptations on the Tabucol algorithm and in his existing implementation in C language, developed by J. Culberson at the University of Alberta, Edmonton (Culberson 2004), in order to allow the solution of the problem considering the additional constraints 1 to 8, described above. The Tabucol algorithm adapted is denoted by Modified Tabucol (MT):

- the σ color classes C_k, $k = 1,\ldots,\sigma$, are grouped in a set of λ ordered μ-tuples $L_r = (C_{k_{r(1)}},\ldots,C_{k_{r(\mu)}})$ called **color groups**, $r = 1,\ldots,\lambda$ and $\sigma = \lambda \cdot \mu$.
- a vertex v^{ρ}_{ij} is in the color group L_r, if this vertex is in a color class $C_k \in L_r$.
- a $\alpha \times \beta$ maximum number of vertices per group matrix $\bar{U} = [\bar{u}_{ij}]$ is defined, corresponding to the $\alpha \times \beta$ daily lessons hours limits matrix $U = [u_{ij}]$, where $\bar{u}_{ij} = u_{ij}$, $0 \le \bar{u}_{ij} \le 2$ and $\bar{u}_{ij}$ is equal to the maximum number of vertices v^{ρ}_{ij} in the same color group.
- a $\alpha \times \beta$ minimum number of consecutive vertices matrix $\bar{S} = [\bar{s}_{ij}]$ is defined, corresponding to the $\alpha \times \beta$ double lessons matrix $S = [s_{ij}]$, where $\bar{s}_{ij} = s_{ij}$, $\bar{s}_{ij} \ge 0$ and $\bar{s}_{ij}$ is equal to the minimum required number of consecutive vertices v^{ρ}_{ij}. Consecutive vertices are defined as two vertices in consecutive color classes of the same color group.
- the objective function is defined by $\bar{f}(c) = \omega \cdot \bar{f}_1(c) + \delta \cdot \bar{f}_2(c) + \pi \cdot \bar{f}_3(c)$, with $\bar{f}_1(c)$, $\bar{f}_2(c)$ and $\bar{f}_3(c)$ components, respectively, corresponding to $f_1(Q)$, $f_2(Q)$ and $f_3(Q)$ components of $f(Q)$, so that:
  - $\bar{f}_1(c) = \sum_{k=1}^{\sigma} |E(C_k)|$, where $E(C_k)$ is the set of edges with both terminal vertices in the color class C_k, $k = 1,\ldots,\sigma$;
  - $\bar{f}_2(c) = \sum_{r=1}^{\lambda} |V(L_r)|$, where $|V(L_r)|$ is the number of vertices v^{ρ}_{ij} in the color group L_r that exceeds the vertices per group limit $\bar{u}_{ij}$;
  - $\bar{f}_3(c) = \sum_{i=1}^{\alpha} (\theta \cdot \bar{g}_i + \varphi \cdot \bar{v}_i + \psi \cdot \bar{l}_i)$, where:

- $\bar{g}_i = \sum_{r=1}^{\lambda} |C_{(L_r)}^{(i)}|$, where $|C_{(L_r)}^{(i)}|$ is the number of color classes of a color group L_r with no vertex $v \in V_i$ and that are positioned in ordered μ-tuple between color classes with at least one vertex $v \in V_i$;
- $\bar{v}_i$ is the number of color groups with at least one vertex $v \in V_i$;
- $\bar{l}_i$ is equal to the non-negative difference between the required value $\sum_{j=1}^{\beta} \bar{s}_{ij}$ and the number of vertices pairs $v, w \in V_i$, such that v and w are consecutive vertices, that is, vertices in consecutive color classes C_{k_r}, C_{k_r+1} of the same color group.

Thus, the function $\bar{f}(c) = \bar{f}_1(c)$, replaced by $\bar{f}(c) = \omega \cdot \bar{f}_1(c) + \delta \cdot \bar{f}_2(c) + \pi \cdot \bar{f}_3(c)$ is the objective function of the graph coloring problem related to CTTP and the concepts described above are incorporated in the Tabucol algorithm. Therefore, a σ-coloring of the adjunct graph G'', obtained by the algorithm, corresponds to a CTTP feasible solution, considering all its hard and soft constraints.

4 Computational Results and Conclusions

The correspondence between the CTTP and GCP is known in the literature (Neufeld and Tartar 1974, 1975), but normally the CTTP contains only basic constraints. In the work presented here, an extended version of this correlation and adaptations on Tabucol algorithm are proposed, as well as modifications on an existing implementation in C language were carried out in order to contemplate the CTTP with additional constraints. The MT algorithm presented here is evaluated in the experiment environment described as follows.

Computational results were obtained for five instances related to the problem, extracted from real timetabling of Brazilians high schools and also two artificial instances with different number of teachers, students, unavailable periods, daily lessons per class with the same teacher and lessons assigned to two consecutive periods (Santos et al 2007; Souza et al 2003). The MT algorithm was tested on a Pentium II 450 MHz PC with 128 MB RAM running CentOS 4.3 Linux operating system and GCC 3.4.5 compiler. The following weights were used in the objective function: $\omega = 100$, $\delta = 30$, $\pi = 1$, $\theta = 3$, $\varphi = 9$ and $\psi = 1$ (Santos et al 2007; Souza et al 2003). For the generation of the initial solution of the MT algorithm, a number of four of the greedy algorithms as described in (Culberson 2004) were used.

The computational results of the MT algorithm for the seven instances were obtained by combining the greedy algorithms used to generate the initial solution and the following parameters (Bello 2007): maximum number of neighbours {500, 800, 900, 1000}, minimum number of neighbours {2, 3, 13, 19, 20} to be generated and tabu list size {7, 10}, which produced the best results in initial tests. The executions were carried out with a fixed number of iterations (30,000) for each instance. The initialization was done with a single seed value for the pseudo-random numbers (seed = 1) for all instances, since this value provided the best overall results in these initial tests. The effect of this approach is verified by comparing the results for the GCP (using the MT algorithm) with those available in the literature for the CTTP (Grasp Tabu Search (GTS) (Souza et al 2003) and Tabu Search (TS) (Santos et al 2007)), but without the corresponding improvement and diversification strategies, since they were not originally implemented in the Tabucol algorithm. The results for the TS are recalculated with the same number of iterations and the same seed for pseudo-random numbers used in the MT and performed in the same machine. For TS, in this experiment the search is executed in the entire neighborhood and the tabu list size was set as proposed in (Santos et al 2007). For the GTS, the results presented in (Souza et al 2003) were reproduced here,

Table 1 Results for GTS, TS and MT algorithms

instance	GTS	TS	MT	TS (ms)	MT (ms)	$\Delta time(\%)$
1	-	204	202	12.96	5.17	150.8
2	368	344	347	36.46	6.82	434.3
3	491	448	436	39.31	7.65	413.9
4	749	671	683	62.64	8.63	625.7
5	831	795	808	106.45	8.92	1093.6
6	847	783	872	115.69	8.83	1210.2
7	1164	1066	1117	185.53	12.63	1368.8

in which instance 1 was not included. The purpose was to provide an additional source for comparison of the results, even without the same standardization used in the comparison of MT with TS. To attenuate this aspect, it was used a machine with similar configuration to the one used in (Souza et al 2003).

The Table 1 shows the average best solution costs (second column) obtained by GTS and the best solution costs (third and fourth columns) obtained respectively by TS and MT when applied to the seven instances (first column) with the parameters described above and the average execution times in milliseconds of TS and MT algorithms (fifth and sixth columns) to achieve the respective best solutions. The last column shows the percentage variations observed in the times spent by both algorithms (TS and MT). The MT algorithm achieved good solutions for the Class/Teacher Timetabling Problem in a very short CPU time when compared with TS.

MT is competitive with respect to the quality of the solutions, as it achieved the optimal solution for the instance 1 (Santos 2007) and the average results are only 3.57% greater than those obtained by TS. When comparing the CPU time, the proposed algorithm, MT, obtained an outstanding performance, since the average CPU time are 853.11% less than the other algorithms. Moreover, it is observed that MT is robust with respect to the size of the instance. The standard deviation of CPU time required by all instances is 2.31, whereas the standard deviation of the TS CPU time is 59.8. As additional source of comparison, the solution costs of GTS algorithm, found in literature for the instances of this problem, are used. The approach proposed presents interesting contributions to the CTTP, and allows future research such as: MT algorithm with long-term memories strategies (diversification and intensification); adaptations to other graph coloring algorithms to consider additional restrictions to CTTP and generalize the process described to other timetable problems.

Acknowledgements We thank Maria Cristina Silva Boeres for her assistance during the English translation of this document.

References

Bello GS (2007) An approach for the timetabling problem subject to constraints using graph coloring (in portuguese). Master's thesis, Universidade Federal do Espírito Santo, Brazil

Culberson J (2004) Graph coloring programs. Joseph Culberson's Graph Coloring Page. http://web.cs.ualberta.ca/~joe/Coloring/index.html

Hertz A, de Werra D (1987) Using tabu search techniques for graph coloring. Computing 39:345–351

Neufeld G, Tartar J (1974) Graph coloring conditions for the existence of solutions to the timetable problem. Communications of the ACM 17(8):450–453

Neufeld G, Tartar J (1975) Generalized graph colorations. SIAM Journal of Applied Mathematics 29(1):91–98

Santos HG (2007) Formulations and algorithms for the scholar timetabling problem (in portuguese). PhD thesis, Universidade Federal Fluminense, Brazil

Santos HG, Ochi LS, Souza M (2007) A tabu search heuristic with efficient diversification strategies for the class/teacher timetabling problem. ACM Journal of Experimental Algorithmics To appear

Souza M (2000) Timetabling in high-schools:approximantion by metaheuristics (in portuguese). PhD thesis, Universidade Federal do Rio de Janeiro (COPPE/UFRJ), Brazil

Souza M, Ochi L, Maculan N (2003) A grasp-tabu search algorithm for solving school timetabling problems. In: Resende M, Souza J (eds) Metaheuristics: Computer Decision-Making, Kluwer Academic Publishers, Boston, pp 659–672