Uma Proposta de Solução para Problemas de Horário Educacional utilizando Busca Dispersa e Reconexão por Caminhos

Morgana Spindler[1], **Leonardo D.Chiwiacowsky**[1]

[1]Programa de Pós-Graduação em Computação Aplicada (PIPCA) - UNISINOS
São Leopoldo – RS – Brasil

morganaspindler@gmail.com, ldchiwiacowsky@unisinos.br

Abstract. *This paper presents a model for solving the problem of constructing the timetables of courses in educational institutions, known in scientific circles as Curriculum Based Course Timetabling. This problem involves the construction of the weekly schedule that indicate the time periods in which each discipline of these courses should take place in compliance with the specifications of the curriculum and allocating teachers and classrooms. The solution technique used is an algorithm based on the Scatter Search metaheuristic that uses the Path Relinking heuristic as a strategy for intensifying the search for optimal solution. The objective of this study is to document the performance of these techniques in solving the problem described, thus generating a database of results.*

Resumo. *Este trabalho apresenta um modelo de solução para o problema de construção das grades de horário de cursos em instituições de ensino, conhecido no meio científico como Programação de Horário de Cursos Baseada em Currículos. Este problema consiste na construção das grades de horários que indicam em quais períodos semanais cada disciplina destes cursos deverá ocorrer, respeitando as especificações dos currículos e alocando professores e salas. A técnica de solução utilizada é um algoritmo baseado na metaheurística Busca Dispersa que utiliza a heurística de Reconexão por Caminhos como estratégia de intensificação de busca da solução ótima. O objetivo deste trabalho é registrar o desempenho destas técnicas na solução do problema descrito, gerando assim, uma base de dados de resultados.*

1. Introdução

Ao longo dos últimos 40 anos, a comunidade científica tem se esforçado para buscar uma solução computacional que auxilie em um antigo, longo e duro processo: a alocação de recursos sob restrições. Um dos problemas típicos desta natureza é o Problema de Programação de Horários Educacionais. O ajuste de interesses entre disponibilidades de salas, professores e alunos ao longo de determinados períodos da semana, é tarefa árdua e exige tempo. Contudo, é essencial e periódica em instituições de ensino, seja semestral ou anualmente, já que todas as atividades são pautadas sobre o período letivo. Esse problema tem sido objeto de pesquisa por parte da comunidade científica, tanto devido ao aumento da sua complexidade, quanto pela busca de melhores caminhos para solucioná-lo.

Segundo Cooper [Cooper and Kingston 1995], o Problema de Programação de Horários Educacionais pode ser modelado como um problema matemático de Otimização Combinatória de complexidade NP-completo, mundialmente referido como *Timetabling Problem* e de grande relevância na área de Pesquisa Operacional. Por sua vez, Wren [Wren 1996] define o problema de *Timetabling* como “a alocação, sujeita a restrições, de recursos a objetos colocados no espaço e no tempo, de modo a satisfazer, tanto quanto possível, um conjunto de objetivos desejáveis".

Dentre as diversas variações do Problema de Programação de Horários Educacionais, este trabalho aborda o problema de Programação de Horário de Cursos Baseada em Currículos. Nesta variação do problema de Horário Educacional, deve-se realizar o agendamento de vários cursos em um conjunto de períodos semanal, considerando também a alocação de um número limitado de salas, sendo que os conflitos entre cursos são determinados de acordo com o currículo dos cursos e não de acordo com as matrículas efetuadas. Neste caso, as turmas representam os módulos curriculares.

Neste trabalho, um método de solução que combina a metaheurística Busca Dispersa [Glover et al. 2003], ou *Scatter Search*, com a heurística Reconexão por Caminhos [Glover 1999] é aplicado para solução do problema de Programação de Horário de Cursos Baseada em Currículos. O trabalho tem como objetivo mostrar a viabilidade do emprego destas técnicas para a solução do referido problema, já que na literatura, o autor não encontrou trabalhos que reportem resultados utilizando esta abordagem.

2. Formulação do problema

O estudo de caso foi realizado considerando os problemas de Programação de Horário de Cursos Baseada em Currículos propostos pela *International Timetabling Competition* (ITC).

Na *International Timetabling Competition* (ITC), o problema de Programação de Horário de Cursos Baseada em Currículos, consiste na programação semanal das aulas para diversos cursos universitários, dado um determinado número de salas e períodos de tempo, onde os conflitos entre os cursos são definidos de acordo com os currículos definidos pela Universidade e não com base nos dados de matrícula de alunos.

Na *International Timetabling Competition*, os pesos correspondentes a cada restrição são definidos por diferentes formulações, onde cada formulação prioriza um tipo ou um conjunto de restrições. A formulação UD1 foi introduzida pelos autores Di Gaspero e Schaerf e considera um número reduzido de restrições [Di Gaspero and Schaerf 2006]. A formulação UD2 é utilizada na edição 2007 da competição [McCollum et al. 2007]. As formulações UD3, UD4 e UD5 [Cesco et al. 2009] priorizam, respectivamente, carga horária dos estudantes, agrupamento de aulas e deslocamento entre aulas.

2.1. Definição do Problema

O problema de Programação de Horário de Cursos Baseada em Currículos consiste em programar as aulas previstas para um período letivo, realizando o agendamento em um conjunto de períodos semanal de vários cursos e alocando um número limitado de professores e salas. A programação das aulas deve atender a um conjunto de requisitos e respeitar as limitações de recursos disponíveis que representam as restrições do problema. Neste caso, os conflitos entre cursos são determinados de acordo com os módulos curriculares dos cursos, representados pelas turmas de alunos, e não de acordo com as matrículas efetuadas.

Nesta formulação, uma *aula* é a alocação de uma *turma*, *professor* e *sala* para a ocorrência de uma *disciplina* em um *horário* (*dia* da semana e *período* do dia).

A carga horária das disciplinas que devem ser agendadas, assim como as turmas que frequentarão as aulas, são dados de entrada do problema. Da mesma forma, as salas disponíveis com seus respectivos recursos também são relacionadas. O professor que deve ministrar cada disciplina é previamente determinado pela instituição. As indisponibilidades de turmas e professores, assim como a exigência de recursos em sala para cada disciplina, é uma informação pré-definida. O número total de aulas que devem ser programadas será definido pelo número de disciplinas $\times$ carga horária de cada disciplina.

2.2. Representação da Solução

O conjunto de aulas programadas por uma solução é representado por uma matriz binária definida na equação (1). Esta matriz representa os períodos de ocorrência das disciplinas quando alocadas a uma respectiva sala. Desta forma, sejam:

- C: conjunto de disciplinas com elementos $c \in C$ e identificados por $c = 1, \dots, \bar{c}$;
- R: conjunto de salas com elementos $r \in R$ e identificados por $r = 1, \dots, \bar{r}$;
- P: conjunto de períodos alocáveis com elementos $p \in P$ e identificados por $p = 1, \dots, \bar{p}$.

Assim, as soluções podem ser representadas pela matriz

$$\mathbf{X}_{c \times r \times p} \ , \tag{1}$$

onde $X_{crp} = 1$ indica que a disciplina c tem uma aula na sala r no período p, e $X_{crp} = 0$, caso contrário.

2.3. Custo da Solução

A importância do cumprimento ou não de uma restrição representa a relevância desta restrição na formulação do problema. A soma ponderada de parcelas associadas a cada restrição violada por uma solução gerada é considerado o custo associado àquela solução. Desta forma, sejam

- V: conjunto de restrições do problema com elementos $v \in V$ e identificados por $v = 1, \dots, \bar{v}$;
- w_v: peso da restrição de avaliação $f_v(\mathbf{X})$ na Função Objetivo (FO);
- $f_v(\mathbf{X})$: função de avaliação da restrição $\mathbf{R}_v$.

A função objetivo, equação (2), minimiza o custo da solução encontrada quando o problema departamental é formulado como um problema de otimização e submetido ao algoritmo de solução:

$$\text{Minimizar} \quad \sum_{v=1}^{\bar{v}} [w_v \times f_v(\mathbf{X})] \ , \tag{2}$$

sujeito às restrições definidas a seguir.

2.4. Restrições

Devido a grande variedade de interesses, sejam organizacionais, pedagógicos ou pessoais, quando da construção de uma grade de horários educacionais, o conjunto de restrições a ser considerado pode apresentar uma grande diferença. Desta forma, as restrições podem ser agrupadas em dois subconjuntos: restrições fortes, cujo cumprimento pela solução encontrada é obrigatório; e restrições fracas, cujo cumprimento pela solução encontrada é desejável.

Desta forma, para definição da formulação matemática das restrições empregadas no problema, foi realizado um estudo bibliográfico onde foi verificada a existência de diferentes formulações. Assim, a formulação matemática aqui apresentada está baseada na combinação de algumas propostas encontradas na literatura [Aladag and Hocaoglu 2007], [Gaspero and Schaerf 2003], [Avella and Vasil'ev 2005], [Alvarez-Valdes et al. 2002].

Inicialmente, são definidas a notação e a definição dos parâmetros utilizados na formulação matemática das restrições, apresentada em seguida. Portanto, sejam:

- C: conjunto de disciplinas com elementos $c \in C$ e identificados por $c = 1, \ldots, \bar{c}$;
- R: conjunto de salas com elementos $r \in R$ e identificados por $r = 1, \ldots, \bar{r}$;
- P: conjunto de períodos alocáveis com elementos $p \in P$ e identificados por $p = 1, \ldots, \bar{p}$;
- D: conjunto de dias da semana com elementos $d \in D$ e identificados por $d = 1, \ldots, \bar{d}$;
- G: conjunto de turmas com elementos $g \in G$ e identificados por $g = 1, \ldots, \bar{g}$;
- V: conjunto de restrições do problema com elementos $v \in V$ e identificados por $v = 1, \ldots, \bar{v}$;
- $\mathbf{DL}_{\bar{c}}$: vetor que define se uma disciplina exige aulas agrupadas ou não, onde $DL_c = 1$ se a disciplina c exige aulas agrupadas, e $DL_c = 0$ caso contrário;
- $\mathbf{CM}_{\bar{c}\times\bar{c}}$: matriz de turmas com disciplinas comuns, onde $CM_{c_1c_2} = 1$ se existe uma turma que inclui as disciplinas c_1 e c_2, e $CM_{c_1c_2} = 0$ caso contrário;
- $\mathbf{TM}_{\bar{c}\times\bar{c}}$: matriz de disciplinas com professores comuns, onde $TM_{c_1c_2} = 1$ se existe um professor que leciona as disciplinas c_1 e c_2, e $TM_{c_1c_2} = 0$ caso contrário;
- $\mathbf{A}_{\bar{c}\times\bar{p}}$: matriz de disponibilidades, onde $A_{cp} = 1$ se as aulas da disciplina c podem ser agendadas no período p, e $A_{cp} = 0$ caso contrário;
- $\mathbf{RS}_{\bar{c}\times\bar{r}}$: matriz de adequação das salas, onde $RS_{cr} = 1$ se as aulas da disciplina c podem ser agendadas na sala r, e $RS_{cr} = 0$ caso contrário;
- C_g: conjunto de disciplinas da turma g;
- B_r: prédio onde a sala r está localizada;
- P_d: conjunto de períodos do dia d;
- l_c: carga horária da disciplina c em número de aulas;
- md_c: mínimo de dias de ocorrência da disciplina c;
- s_c: máximo de alunos matriculados na disciplina c;
- min: mínimo de aulas por dia para cada turma;
- max: máximo de aulas por dia para cada turma;
- cap_r: capacidade da sala r;
- w_v: peso da restrição de avaliação $f_v(\mathbf{X})$ na FO;
- $f_v(\mathbf{X})$: função de avaliação da restrição $\mathbf{R}_v$.

A seguir, são apresentadas as definiçõe de cada uma das restrições.

2.4.1. Restrição *Aulas* [R01]

Esta restrição determina que a carga horária de cada uma das disciplinas deve ser atendida.

$$\sum_{r\in R} \sum_{p\in P} X_{crp} = l_c, \quad \forall\, c \in C. \tag{3}$$

2.4.2. Restrição *Conflitos* [R02]

Esta restrição determina que aulas de disciplinas do mesmo currículo (ou turma) devem ser agendadas em horários diferentes, assim como aulas ministradas pelo mesmo professor, devem ser agendadas em horários diferentes.

$$(X_{c_1rp} + X_{c_2rp})\, CM_{c_1c_2} \leq CM_{c_1c_2}, \quad \forall\, c_1,\ c_2 \in C, \quad \forall\, p \in P, \quad \forall\, r \in R; \tag{4}$$

$$(X_{c_1rp} + X_{c_2rp})\, TM_{c_1c_2} \leq TM_{c_1c_2}, \quad \forall\, c_1,\ c_2 \in C, \quad \forall\, p \in P, \quad \forall\, r \in R. \tag{5}$$

2.4.3. Restrição *Ocupação das salas* [R03]

Esta restrição determina que duas aulas não podem ser agendadas na mesma sala e no mesmo horário.

$$\sum_{c \in C} X_{crp} \leq 1, \quad \forall\, r \in R, \quad \forall\, p \in P. \tag{6}$$

2.4.4. Restrição *Disponibilidade* [R04]

Esta restrição determina que se o professor da disciplina não está disponível para um determinado horário, nenhuma aula da disciplina pode ser agendada nesse período. Da mesma forma, existem horários que determinadas disciplinas não podem ser agendadas, por exemplo, em função do turno de ocorrência do curso.

$$X_{crp} \leq A_{cp}, \quad \forall\, c \in C, \quad \forall\, r \in R, \quad \forall\, p \in P. \tag{7}$$

2.4.5. Restrição *Capacidade das salas* [R05]

Esta restrição determina que para cada disciplina, o número de alunos que frequentam as aulas deve ser menor, ou igual, ao número de lugares existentes na sala alocada.

$$X_{crp} \times s_c \leq cap_r, \quad \forall\, c \in C, \quad \forall\, r \in R, \quad \forall\, p \in P. \tag{8}$$

2.4.6. Restrição *Mínimo de dias de aula* [R06]

Esta restrição determina que as aulas de cada disciplina devem ser distribuídas em um determinado número mínimo de dias.

$$\sum_{d \in D} \left[\mathcal{H}\Big(\sum_{r \in R} \sum_{p \in P_d} X_{crp}\Big) \right] \geq md_c, \quad \forall\, c \in C; \tag{9}$$

onde a notação $\mathcal{H}(\cdot)$, utilizada na formulação desta restrição mas também das restrições [R07] e [R09], indica a função Degrau Unitário ou Função de *Heaviside* discreta [Kanwal 1998].

2.4.7. Restrição *Aulas isoladas* [R07]

Esta restrição (compactação de currículo 1) determina que as aulas pertencentes a um currículo (turma) devem ser adjacentes, ou seja, em períodos consecutivos de um mesmo dia.

$$\sum_{c \in C_g} \sum_{p=p'-1}^{p'+1} \sum_{r \in R} X_{crp}\, \mathcal{H}\Big(\sum_{c \in C_g} \sum_{p \in P_d} \sum_{r \in R} X_{crp}\Big) \geq 2\, \mathcal{H}\Big(\sum_{c \in C_g} \sum_{p \in P_d} \sum_{r \in R} X_{crp}\Big), \quad \forall\, d \in D, \forall\, g \in G, \forall\, p' \in P_d. \tag{10}$$

2.4.8. Restrição *Janelas* [R08]

Esta restrição (compactação de currículo 2) determina que as aulas pertencentes a um currículo (turma) não devem ter janelas de tempo entre elas no mesmo dia, isto é, períodos sem agendamentos.

$$(X_{c''r''p''} \times p'') - (X_{c'r'p'} \times \ p') < \sum_{c \in C_g} \sum_{p=p'}^{p''} \sum_{r \in R} X_{crp},$$

$$\forall\, c', c'' \in C_g, \quad \forall\, r', r'' \in R, \quad \forall\, p', p'' \in P_d, \quad \forall\, X_{c'r'p'} = 1,$$
$$\forall\, X_{c''r''p''} = 1, \quad \forall\, p \geq p', \quad \forall\, p'' > p', \quad \forall\, d \in D, \quad \forall\, g \in G. \tag{11}$$

2.4.9. Restrição *Manutenção de sala* [R09]

Esta restrição determina que todas as aulas de uma disciplina devem ocorrer na mesma sala.

$$\sum_{p \in P} X_{crp} \left[\mathcal{H}(\sum_{p \in P} X_{crp}) \right] = l_c\, \mathcal{H}(\sum_{p \in P} X_{crp}), \quad \forall\, c \in C, \quad \forall\, r \in R. \tag{12}$$

2.4.10. Restrição *Mínimo e máximo de aulas por dia* [R10]

Esta restrição determina que, para cada turma, o número de aulas diárias deve estar dentro de um determinado intervalo, indicando o número mínimo e máximo de aulas.

$$min \leq \sum_{c \in C_g} \sum_{p \in P_d} \sum_{r \in R} X_{crp} \leq max, \quad \forall\, g \in G, \quad \forall\, d \in D. \tag{13}$$

2.4.11. Restrição *Deslocamentos* [R11]

Esta restrição determina que os alunos devem ter tempo para se deslocar de um prédio para outro entre duas aulas. Duas aulas da mesma turma, em períodos adjacentes no mesmo dia, devem ocorrer em salas localizadas no mesmo edifício.

$$(X_{c'r'p'} \times p') - (X_{crp} \times p) > 1,$$

$$\forall\, c, c' \in C_g, \quad \forall\, r, r' \in R \text{ tal que } B_r \neq B_{r'}, \quad \forall\, p, p' \in P_d,$$
$$\forall\, X_{crp} = 1, \quad \forall\, X_{c'r'p'} = 1, \quad \forall\, p' > p, \quad \forall\, d \in D, \quad \forall\, g \in G. \tag{14}$$

2.4.12. Restrição *Adequação de salas* [R12]

Esta restrição garante a adequação da sala, visto que algumas salas podem não ser adequadas para uma determinada disciplina por causa da ausência de recursos necessários (ginásio de esportes, projetor, computadores, etc).

$$X_{crp} \leq RS_{cr}, \quad \forall\, c \in C, \quad \forall\, p \in P, \quad \forall\, r \in R. \tag{15}$$

2.4.13. Restrição *Agrupamento de aulas* [R13]

Esta restrição determina que algumas disciplinas exigem que as aulas agendadas no mesmo dia sejam agrupadas. Duas aulas estão agrupadas se são realizadas na mesma sala e seus horários de realização são adjacentes.

$$[(X_{cr'p''} \times p'') - (X_{cr'p'} \times p')] \; DL_c < \sum_{r \in R} \sum_{p=p'}^{p''} X_{crp} \; DL_c,$$

$$\forall\, c \in C_g, \quad \forall\, r' \in R, \quad \forall\, p', p'' \in P_d, \quad \forall\, X_{crp} = 1,$$
$$\forall\, X_{crp'} = 1, \quad \forall\, p'' > p', \quad \forall\, p \geq p', \forall\, d \in D, \quad \forall\, g \in G. \qquad (16)$$

3. Método de solução

Para solução do problema apresentado, foi desenvolvida uma ferramenta baseada no uso da metaheurística Busca Dispersa com uso da heurística Reconexão por Caminhos.

A Busca Dispersa ou *Scatter Search* (SS), é um método que foi formalmente introduzido por Glover [Glover 1998] em um estudo heurístico para a resolução de problemas de Programação Linear Inteira.

Em comum com outros métodos populacionais, a Busca Dispersa opera sobre um conjunto de soluções, ao invés de uma única solução em cada iteração, e emprega procedimentos para combinar essas soluções com o intuito de gerar novas soluções. Por outro lado, diferentemente desses mesmos métodos, ao invés de explorar soluções de forma aleatória, como nos Algoritmos Genéticos, a Busca Dispersa explora extensivamente regiões pré-determinadas em cada iteração. A Busca Dispersa opera sobre um Conjunto de Soluções Referência ($RefSet$) para gerar novas soluções a partir de combinações lineares ponderadas das soluções pertencentes a subconjuntos estruturados do $RefSet$ [Glover 1998].

O $RefSet$ é composto pelas melhores soluções encontradas. O conceito de melhores neste caso não considera apenas a característica de estar entre os melhores valores da função objetivo (FO), mas leva também em consideração a diversidade da solução em relação as outras pertencentes ao $RefSet$ [Glover et al. 2003].

De acordo com Glover [Glover et al. 2003], o algoritmo consiste basicamente de cinco elementos: Método de geração de soluções diversas, Método de melhoria, Método de atualização do conjunto de referência, Método de geração de subconjuntos e Método de combinação das soluções.

Implementações mais recentes da metaheurística Busca Dispersa, ou *Scatter Search*, foram propostas por Glover e colaboradores [Glover et al. 2003], onde é examinado o seu uso conjunto com a heurística Reconexão por Caminhos (*Path Relinking*) na solução de diferentes problemas de otimização combinatória.

A técnica Reconexão por Caminhos, Religação de Caminhos ou *Path Relinking* (PR), foi proposta por Glover [Glover 1999] como uma estratégia de intensificação, para explorar trajetórias que conectavam soluções elite obtidas pelos métodos Busca Tabu ou Busca Dispersa.

Esta busca por soluções de melhor qualidade consiste em gerar e explorar caminhos no espaço de soluções partindo de uma ou mais soluções elite e levando a outras soluções elite. Para tal finalidade, são selecionados movimentos que introduzem atributos das soluções guia na solução corrente. Desse modo, a Reconexão por Caminhos pode ser vista como uma estratégia que tem por objetivo incorporar atributos de soluções de boa qualidade, favorecendo a seleção de movimentos que as contenham.

O método de Reconexão por Caminhos inicia computando a diferença de simetria entre duas soluções s_I e s_G, resultando em um conjunto de movimentos que deve ser aplicado a uma delas, dita solução inicial, para transformar-se na outra, dita solução guia. A partir da solução inicial, o melhor movimento ainda não executado é aplicado à solução corrente, até que a solução guia seja atingida. O atributo a ser alterado na solução inicial é definido aleatoriamente, dentre aqueles possíveis, podendo ser a modificação de uma sala, a alteração de um horário ou a troca de disciplina. A melhor solução encontrada ao longo dessa trajetória é considerada como candidata à inserção no conjunto elite e a melhor solução já encontrada é atualizada. Considere o caminho que liga as soluções s_I e s_G, onde s^i representa as diferentes soluções obtidas pela seqüência de aplicação dos movimentos, tal que

$$s_I = s^1,\, s^2, \ldots,\, s^r = s_G\,;$$

onde a solução $s^{k+1} \in N_{PR}(s^k)$ é obtida pela aplicação de um movimento e N_{PR} é a vizinhança considerada.

3.1. Algoritmo Implementado

O processo de otimização da metaheurística Busca Dispersa inicia com o emprego do método de geração de soluções diversas, que tem por objetivo gerar uma população P de soluções diversas e com tamanho $PSize$. Tal método tem sua aplicação considerando apenas a restrição de que todas as aulas sejam programadas, desconsiderando violações às demais restrições do problema.

Uma vez geradas, as soluções são submetidas ao método de melhoria, que visa eliminar violações às restrições fortes, as quais inviabilizam uma solução, e também diminuir as violações às restrições fracas, que influenciam na otimalidade da solução.

A partir daí, através do emprego do método de criação do conjunto de referência, um conjunto menor, de tamanho b (geralmente com tamanho igual a 10% do tamanho de P) é gerado com as b_1 melhores soluções (aquelas com menor custo) e as b_2 soluções mais diversas (aquelas com um maior grau de diversidade em relação às melhores) de P. Este conjunto de soluções de referência é denotado por $RefSet$.

O conceito de melhores soluções está relacionado ao custo destas soluções. Uma solução é melhor que outra se o seu custo é menor. O conceito de diversidade está relacionado ao número de diferenças entre uma solução e outra, ou em relação a um conjunto de soluções, onde cada atributo divergente é contabilizado como uma diferença.

Após esta etapa, inicia-se um processo iterativo onde, a cada iteração, as soluções presentes no $RefSet$ são combinadas (Reconexão por caminhos) e melhoradas, até que um critério de parada seja satisfeito. O critério de parada consiste no atendimento a uma das três hipóteses seguintes: (i) quando for encontrada uma solução viável de custo zero; (ii) quando o processo iterativo demonstrar que nenhuma nova solução, de melhor qualidade (ou diferente e de mesma qualidade), pode ser incorporada ao $RefSet$, quando oriunda da combinação e/ou melhoria das soluções anteriores e; (iii) quando um número máximo de iterações é atingido, evitando que a aplicação fique indefinidamente em execução.

Fazendo parte do processo iterativo, o método de geração de subconjuntos e o método de combinação das soluções são responsáveis pela criação do conjunto $Pool$, formado pelas soluções combinadas do $RefSet$. Antes de substituirem as piores e menos diversas soluções de $RefSet$, as soluções presentes em $Pool$ são submetidas ao método de melhoria e posteriormente avaliadas. Esta etapa é realizada pelo método de atualização do conjunto de referência, que tem como objetivo melhorar a qualidade das soluções, através da eliminação das violações às restrições fortes e diminuição das violações às restrições fracas, tendo como resultado soluções viáveis de melhor qualidade.

A avaliação de uma solução significa contabilizar o número de violações a cada tipo de restrição (conforme definidas na Seção 2.4) que esta solução apresenta e, assim, determinar o seu custo, conforme a equação (2). A solução final do problema será a melhor solução de $RefSet$, ou seja, aquela solução que apresentar o menor custo após o encerramento do processo iterativo.

4. Validação e resultados

O arquivo com os resultados da programação de horários, no caso das instâncias da *International Timetabling Competition* , pode ser submetido a um validador da competição [1]. Portanto, para validação do modelo, foi selecionada a instância *comp01* da *International Timetabling Competition*.

Foram realizados três testes com a instância *comp01*, conforme as definições de pesos das formulações UD2, UD3 e UD4. Os testes foram denominados ITC_{11}, ITC_{12} e ITC_{13}, respectivamente, e seus resultados são apresentados nas tabelas 2–4.

As médias e o desvio padrão, para cada caso de teste, foram calculados com base em cinco simulações, sendo os parâmetros utilizados na metaheurística Busca Dispersa definidos na tabela 1. O tempo médio de processamento de cada simulação foi de 2 horas.

Tabela 1: Parâmetros dos testes *ITC*.

$Psize$	$b1$	$b2$	ciclos de melhoria de P	ciclos de melhoria de $Pool$	máximo de vizinhanças	critério de parada (número de ciclos)
500	5	5	10	100	6	200

Tabela 2: Testes ITC_{11}.

nome: ITC_{11} instância: comp01 pesos: UD2													
restrições	R01	R02	R03	R04	R05	R06	R07	R08	R09	R10	R11	R12	R13	custo
pesos	100	100	100	100	1	5	2	0	1	0	0	0	0	-
melhor ITC	0	0	0	0	4	0	0	-	1	-	-	-	-	5
melhor SS	0	0	0	0	31	0	3	24	12	2	35	21	13	49
média	0	0	0	0	37,6	0	6,2	25,2	16,8	6,8	50,4	22	29,6	66,8
desvio padrão	0	0	0	0	10,96	0	1,78	3,03	3,27	2,94	11,65	4,84	5,41	15,49

A linha melhor ITC representa os valores de comparação, ou seja, o custo relativo a cada restrição (indicadas pelas respectivas colunas) do melhor resultado obtido para a formulação UD2 e registrado na competição [2].

Para o teste ITC_{11}, a simulação de melhor resultado forneceu uma função objetivo com custo 49. A média de custo, considerando-se todos os testes foi de 66,8 com desvio padrão de 15,49. Vale salientar que, apesar do custo final devido à restrição R05 nesta solução

[1] http://tabu.diegm.uniud.it/ctt/index.php?page=valid
[2] http://tabu.diegm.uniud.it/ctt/index.php?page=rankings&see=best

ser 31, isto representa apenas 7 salas com alocação superior à capacidade permitida. Da mesma forma, o custo final 6 para a restrição R07 representa apenas 3 aulas isoladas.

Portanto, a restrição R09, correspondente a ocorrência de todas as aulas de uma disciplina na mesma sala, foi a restrição que apresentou maior número de violações, ou seja, 12.

Os resultados do teste ITC_{11} demonstram que o método foi capaz de melhorar 96,80%, em média, o custo da solução final em relação à solução inicial, com desvio padrão de 0,88%. Em média, a solução final é obtida no ciclo 125, com desvio padrão de 25,45 ciclos.

Tabela 3: Testes ITC_{12}.

nome: ITC_{12} instância: comp01 pesos: UD3														
restrições	R01	R02	R03	R04	R05	R06	R07	R08	R09	R10	R11	R12	R13	custo
pesos	100	100	100	100	1	0	0	4	3	2	0	0	0	-
melhor ITC	0	0	0	0	4	-	-	0	0	4	-	-	-	8
melhor SS	0	0	0	0	29	16	3	4	13	5	56	29	18	94
média violações	0	0	0	0	40,4	12	5,8	5,2	14,4	6,4	43,2	22,6	22,6	117,2
desvio padrão	0	0	0	0	13,83	2,34	1,64	1,30	1,67	1,34	8,72	3,97	3,84	17,76

Para o teste ITC_{12}, a simulação de melhor resultado forneceu uma função objetivo com custo 94. A média de custo, considerando-se todos os testes foi de 117,2 com desvio padrão de 17,76.

Os resultados do teste ITC_{12} demonstram que o método foi capaz de melhorar 94,70%, em média, o custo da solução final em relação à solução inicial, com desvio padrão de 0,74%. Em média, a solução final é obtida no ciclo 145, com desvio padrão de 20,42 ciclos.

Tabela 4: Testes ITC_{13}.

nome: ITC_{13} instância: comp01 pesos: UD4														
restrições	R01	R02	R03	R04	R05	R06	R07	R08	R09	R10	R11	R12	R13	custo
pesos	100	100	100	100	1	1	0	1	0	1	0	100	1	-
melhor ITC	0	0	0	0	4	1	-	0	-	1	-	0	0	6
melhor SS	0	0	0	0	52	2	34	14	14	7	48	0	14	83
média violações	0	0	0	0	55,6	2	34,4	14	14,6	7,2	49,8	0	16,8	95,6
desvio padrão	0	0	0	0	7,30	0,70	3,28	0,70	1,14	0,44	2,94	0	3,27	9,34

Para o teste ITC_{13}, a simulação de melhor resultado forneceu uma função objetivo com custo 83. A média de custo, considerando-se todos os testes foi de 95,6 com desvio padrão de 9,34.

Os resultados do teste ITC_{13} demonstram que o método foi capaz de melhorar 95,69%, em média, o custo da solução final em relação à solução inicial, com desvio padrão de 0,66%. Em média, a solução final é obtida no ciclo 147, com desvio padrão de 22,12 ciclos.

Em todas as simulações, o método obtém melhoras consecutivas no custo das soluções, em média, até o ciclo 30. A partir deste ponto do processo iterativo de solução, as melhoras ocorrem, em média, a cada 8 ciclos, até que o ótimo local seja encontrado.

A Figura 1, correspondente ao teste de melhor resultado (ITC_{11}), apresenta esta evolução. O gráfico mostra uma melhora acentuada no custo das soluções até o ciclo 25. Do ciclo 25 ao 96 as melhoras ocorrem, na média, a cada 8 ciclos. A solução final é obtida no ciclo 96. O custo da solução final é 97,24% melhor do que a melhor solução inicial.

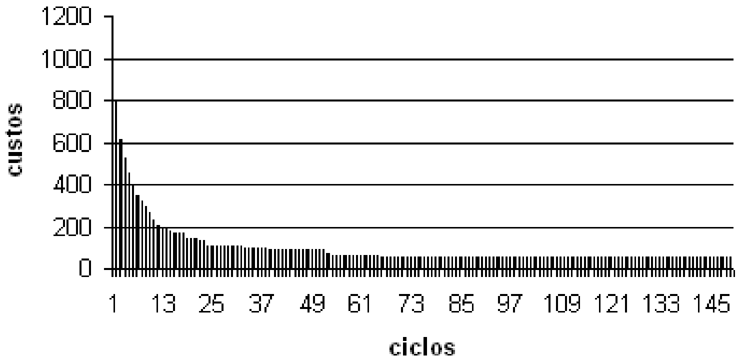

Figura 1: Evolução da melhor solução em $RefSet$.

A implementação do modelo e a sua execução sobre dados da *International Timetabling Competition* proporcionaram a comprovação de sua aplicabilidade, porém, em nenhuma das simulações o método foi capaz de encontrar o ótimo global, ficando preso em soluções que aparentemente representam ótimos locais.

5. Conclusões

O objetivo deste trabalho foi registrar o desempenho de um algoritmo baseado na metaheurística Busca Dispersa e na heurística Reconexão por Caminhos, quando aplicado na solução de instâncias de problemas de programação de cursos pós matrícula, abordados na *International Timetabling Competition*.

Uma das principais dificuldades encontrada no desenvolvimento desta pesquisa foi a falta de referências na literatura para a implementação dos métodos Busca Dispersa e Reconexão por Caminhos para o problema de Programação de Horário de Cursos Baseada em Currículos. Os autores encontraram apenas um trabalho atestando a aplicação da metaheurística Busca Dispersa para o subproblema da classe de Problemas de Horários denominado Problema de Programação de Exames [Mansour et al. 2009]. Portanto, o registro de implementação desta técnica de solução é uma das contribuições deste trabalho.

A formulação das restrições apresentada, é aplicável ao problema de Programação de Horário de Cursos Baseada em Currículos da *International Timetabling Competition* e os autores não encontraram nenhuma publicação que apresentasse a formulação matemática completa das restrições consideradas neste problema. Registra-se neste fato outra contribuição deste trabalho. A escolha das instâncias de problemas da *International Timetabling Competition* deveu-se à necessidade de comparação dos resultados obtidos e, consequentemente, para validação da eficácia do método para o tipo de problema abordado. A implementação do modelo e sua execução sobre dados da *International Timetabling Competition* proporcionaram a comprovação de sua aplicabilidade, mesmo que com resultados inferiores aos registrados anteriormente por outros métodos de solução.

Referências

Aladag, C. H. and Hocaoglu, G. (2007). A tabu search algorithm to solve a course timetabling problem. *Hacettepe Journal of Mathematics and Statistics*, 36(1):53 – 64.

Alvarez-Valdes, R., Crespo, E., and Tamarit, J. M. (2002). Design and implementation of a course scheduling system using tabu search. *European Journal of Operational Research*, 137(3):512–523.

Avella, P. and Vasil'ev, I. (2005). A computational study of a cutting plane algorithm for university course timetabling. *Journal of Heuristics*, 8(6):497–514.

Cesco, F., Gaspero, L., and Schaerf, A. (2009). Benchmarking curriculum-based course timetabling: Formulations, data formats, instances, validation, and results. Technical report, University of Undine, Undine, Italy.

Cooper, T. B. and Kingston, J. H. (1995). The complexity of timetable construction problems. In *PATAT 1995*.

Di Gaspero, L. and Schaerf, A. (2006). Neighborhood portfolio approach for local search applied to timetablingproblems. *Journal of Mathematical Modeling and Algorithms*, 5(1):65–89.

Gaspero, L. D. and Schaerf, A. (2003). Multi-neighbourhood local search with application to course timetabling. In *Proceedings of the 4th International Conference on the Practice and Theory of Automated Timetabling (PATAT-2002), number 2740 in Lecture Notes in Computer Science*, pages 262–275. Springer-Verlag.

Glover, F. (1998). A template for scatter search and path relinking. In *Artificial Evolution '97: Selected Papers from the Third European Conference on Artificial Evolution*, pages 13–54, London, UK. Springer-Verlag.

Glover, F. (1999). *Scatter search and path relinking*. McGraw-Hill Ltd., Maidenhead, UK, England.

Glover, F., Laguna, M., and Martí, R. (2003). *Scatter search*. Springer-Verlag Inc., New York, NY, USA.

Kanwal, R. P. (1998). *Generalized Functions: Theory and Technique*. Birkhäuser, Boston, USA.

Mansour, N., Isahakian, V., and Ghalayini, I. (2009). Scatter search technique for exam timetabling. *Applied Intelligence*, pages 13–54.

McCollum, B., Gaspero, L. D., and Schaerf, A. (2007). The second international timetabling competition (ITC-2007): Curriculum-based course timetabling (Track 3). Technical report, University of Undine, Undine, Italy.

Wren, A. (1996). Scheduling, timetabling and rostering - a special relationship? In *Selected papers from the First International Conference on Practice and Theory of Automated Timetabling*, pages 46–75, London, UK. Springer-Verlag.